DOCUMENT RESUME

ED 462 148 PS 030 066

AUTHOR	Karoly, Lynn A.; Kilburn, M. Rebecca; Bigelow, James H.; Caulkins, Jonathan P.; Cannon, Jill S.
TITLE	Assessing Costs and Benefits of Early Childhood Intervention Programs: Overview and Application to the Starting Early Starting Smart Program [with] Executive Summary.
INSTITUTION	Rand Corp., Santa Monica, CA.
SPONS AGENCY	Casey Family Programs, Seattle, WA.; Substance Abuse and Mental Health Services Administration (DHHS/PHS), Rockville, MD.
ISBN	ISBN-0-8330-2973-8; ISBN-0-9708278-1-4; ISBN-0-8330-2974-6; ISBN-0-97082-78-0-6
PUB DATE	2001-00-00
NOTE	202p.; Appended Executive Summary also produced by James R. Chiesa. Administered by the HAMHSA Office on Early Childhood. see PS 030 067.
AVAILABLE FROM	Casey Family Programs, 1300 Dexter Avenue North, Suite 300, Seattle, WA 98109. Tel: 206-282-7300; Fax: 206-378-4619; Web site: http://www.Casey.org.
PUB TYPE	Reports - Evaluative (142)
EDRS PRICE	MF01/PC09 Plus Postage.
DESCRIPTORS	*Cooperative Programs; *Cost Effectiveness; Costs; Early Childhood Education; Early Intervention; *Evaluation Methods; *Evaluation Problems; Family Programs; Health Programs; Health Services; Longitudinal Studies; Models; Program Effectiveness; Program Evaluation; Young Children
IDENTIFIERS	School Based Services

ABSTRACT

Starting Early Starting Smart (SESS) is a knowledge development initiative designed to: (1) create and test a new model for providing mental health and substance abuse prevention and treatment for young children (birth to 7 years) and their families; and (2) inform practitioners and policymakers of successful interventions and promising practices from a multi-year study. In the third year of the first 5-year phase, 12 culturally diverse grantee organizations provide integrated behavioral health services in community-based early childhood settings where families customarily receive services for children. This report identifies the conceptual and methodological issues associated with the analysis of costs and outcomes of early intervention programs in general and makes recommendations regarding the application of these tools for subsequent demonstration studies of SESS. Chapter 1 discusses the results-based accountability movement in social services and argues that knowledge about the relationship between costs and outcomes is useful for program directors but is also important for developing policy approaches at a more general level. Chapter 2 provides a primer on various types of cost and outcome analysis: cost-benefit, cost-effectiveness, and related methods within the broader decision support framework or policy scorecard analysis. Chapter 3 discusses issues in cost and outcome analysis specific to early childhood intervention programs, focusing on measures for benefits and estimation of costs. Chapter 4 reviews the literature on cost and outcome analysis for early childhood intervention programs and provides illustrations of such analyses from the Perry Preschool Program, the Elmira Prenatal/Early Infancy

Reproductions supplied by EDRS are the best that can be made
from the original document.

Project, and the Chicago Child Parent Centers. Chapter 5 applies the concepts to the SESS program, with specific recommendations regarding the evaluation design and implementation of cost and outcome analysis. Chapter 6 summarizes the main findings and presents conclusions. Three appendices include descriptions of the grantees and mission statements of the national collaborators involved in the project. Also appended is the Executive Summary. (Contains 1 figure, 15 tables, and a 77-item bibliography.) (KB)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 462 148

Assessing Costs and Benefits of Early Childhood Intervention Programs:

Overview and Application to the Starting Early Starting Smart Program

[with] Executive Summary

Lynn A. Karoly
M. Rebecca Kilburn
James H. Bigelow
Jonathan P. Caulkins
Jill S. Cannon

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
● Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PS 030 066

BEST COPY AVAILABLE

SESS

STARTING EARLY STARTING SMART

Early

Assessing Costs and Benefits of Early Childhood Intervention Programs

Overview and Application to the Starting Early Starting Smart Program

Lynn A. Karoly
M. Rebecca Kilburn
James H. Bigelow
Jonathan P. Caulkins
Jill S. Cannon

BEST COPY AVAILABLE

Smart

ABOUT STARTING EARLY STARTING SMART

Starting Early Starting Smart (SESS) is a knowledge development initiative designed to

- Create and test a new model for providing integrated behavioral health services (mental health and substance abuse prevention and treatment) for young children (birth to 7 years) and their families; and to
- Inform practitioners and policymakers of successful interventions and promising practices from the multi-year study, which lay a critical foundation for the positive growth and development of very young children.

In October 1997, with initial funding of $30 million, the Substance Abuse and Mental Health Services Administration (SAMHSA) and Casey Family Programs embarked on a precedent-setting public/private collaboration. Twelve culturally diverse grantee organizations were selected. Each provides integrated behavioral health services in community-based early childhood settings—such as Child Care, Head Start and Primary Care Clinics—where young families customarily receive services for children. Critical to this project is the required collaboration among funders, grantees, consumers, and local site service providers. Implicit in the design of this project is sustainability planning for secured longevity of the programs.

The SESS approach informs policy-making for:

- Service system redesign
- Strengthening the home environment
- Using culture as a resource in planning services with families
- Service access and utilization strategies
- Targeting benefits for children
- Working with families from a strengths-based perspective

The Research Design

The 12 grantees, working collaboratively, designed a study whereby integrated behavioral health services are delivered in typical early childhood settings. Each site has an intervention and comparison group, and each site delivers similar targeted, culturally-relevant, interventions for young children and their families. A collaboratively determined set of outcomes has been established to evaluate project effectiveness:

- Access to and use of services
- Social, emotional, and cognitive outcomes for children
- Caregiver-child interaction outcomes
- Family functioning

The goal of the SESS research is to provide rigorous scientific evidence concerning whether children and families participating in SESS programs achieve better access to needed services and better social, emotional, cognitive, and behavioral health outcomes than do the children and families not receiving these services. SESS programs may also generate information about opportunities, practices, and barriers to sought-after outcomes. This information is critical to achieving effective public policies.

SESS Extended

It was clear from the early days of SESS that whatever effects were uncovered, longitudinal extension of the study would be valuable. In 2001, SAMHSA and Casey Family Programs embarked upon an extension phase, which will increase understanding of the impact of early intervention as young children enter preschool and school years, when babies or toddlers are asked to meet escalating emotional and cognitive demands. This longitudinal extension can validate early methods and findings and assess their durability. It is anticipated that this work will include additional data points of a refined instrument set and intervention package with the addition of study questions related to cost and value, and other special studies. Additional future plans include applying and validating early SESS lessons learned, key concepts, components, and principles to new settings that serve families with young children.

Summation

In sum, SESS reflects the growing acknowledgement that it is important to target positive interventions to very young children. The infant and preschool years lay a critical foundation for later growth and development. Second, successful interventions for very young children must meet the multiple behavioral health, physical health, and educational needs of families. Third, integrated behavioral health services must be made more accessible to families with multiple needs, which are difficult to meet in a fragmented service system.

For more information about *Starting Early Starting Smart* and related SAMHSA-Casey products, contact *www.casey.org* or *www.samhsa.gov* (SESS section under construction).

BEST COPY AVAILABLE

Assessing Costs and Benefits of Early Childhood Intervention Programs

Overview and Application to the Starting Early Starting Smart Program

Lynn A. Karoly
M. Rebecca Kilburn
James H. Bigelow
Jonathan P. Caulkins
Jill S. Cannon

Spring 2001

This report would not have been possible without the contributions of staff from the Office on Early Childhood, SAMHSA, U.S. Department of Health and Human Services, the Casey Family Programs, the Starting Early Starting Smart principal investigators, project directors and researchers, and the parent representatives, who helped design and supervise the data collection. The content of this publication does not necessarily reflect the views or policies of the U.S. Department of Health and Human Services, or the Casey Family Programs, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. government. Responsibility for the content of this report, however, rests solely with the named authors.

Library of Congress Cataloging-in-Publication Data

Assessing costs and benefits of early childhood intervention programs : overview and
applications to the Starting Early Starting Smart program / Lynn A. Karoly ... [et al.].
p. cm.
"MR-1336-CFP"—T.p., v. 1.
"MR-1336/1-CFP"—T.p., v. 2.
Includes bibliographical references.
Contents: [v. 1. without special title]—[v. 2] Executive summary.
ISBN 0-9708278-1-4 (v. 1)—ISBN 0-9708278-0-6 (v. 2)—Casey
ISBN 0-8330-2973-8 (v. 1)—ISBN 0-8330-2974-6 (v. 2)—RAND
1. Starting Early Starting Smart (Program) 2. Child welfare—United States—
Evaluation. 3. Early childhood education—United States—Evaluation. 4. Child mental
health services—United States—Evaluation. 5. Children—Drug use—United States—
Evaluation. I. Karoly, Lynn A., 1961– II. Starting Early Starting Smart (Program).

HV741 .A86 2001
362.7'0973—dc21

2001019040

Any or all portions of this document may be reproduced with proper citation: "Source: Karoly, L., Kilburn, R., Bigelow, J. H., Caulkins, J. P., and Cannon, J. S. (2001). *Assessing Costs and Benefits of Early Childhood Intervention Programs: Overview and Applications to the* Starting Early Starting Smart *Program.* Publishers: Seattle: Casey Family Programs; Santa Monica: RAND."

Published 2001 by Casey Family Programs
1300 Dexter Avenue North, Suite 300
Seattle, WA 98109
Telephone: (206) 282-7300; Fax (206) 378-4619;
Internet: www.Casey.org

RAND
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1200 South Hayes Street, Arlington, VA 22202-5050
201 North Craig Street, Suite 102, Pittsburgh, PA 15213
Internet: www.rand.org/
To order from RAND, contact Distribution Services:
Telephone: (310) 451-7002; Fax: (310) 451-6915;
E-mail: order@rand.org

RAND is a nonprofit institution that helps improve
policy and decisionmaking through research and analysis.
RAND® is a registered trademark.

PREFACE

The increased interest in the potential for early childhood intervention programs to save dollars in the long run has focused attention on the potential for cost-benefit and related analyses to aid decisionmakers in their policy choices. The goal of this report is to identify the conceptual and methodological issues associated with the analysis of costs and outcomes of early intervention programs in general and to make recommendations regarding the application of these tools for subsequent demonstration studies of a particular intervention program: *Starting Early Starting Smart* (SESS).

SESS is a public-private collaboration designed to test the effectiveness of integrating behavioral health services within primary care and early childhood service settings for children from birth to age seven. The SESS program is an initiative of the Office on Early Childhood, Substance Abuse and Mental Health Services Administration (SAMHSA), and the Casey Family Programs, along with several other federal sponsors. The program currently operates in 12 sites across the United States and is entering the third year of its first five-year phase. An outcomes evaluation is built into the first phase.

Program sponsors are beginning to plan for a second phase, the design of which they hope will be informed by the first phase. It was during the initiation of this planning process that program sponsors identified a need for cost information to supplement their outcomes information. Recognizing that the literature offered somewhat limited guidance on the specifics of cost considerations in this context, they requested that RAND not only present them with a summary of research bearing on their problem but that we also examine their program and make specific recommendations regarding how cost and outcome analysis could improve their decisionmaking.

This project began with a meeting of cost and outcome analysis experts held in August 2000, convened by RAND on behalf of the Casey Family Programs and the Office on Early Childhood, SAMHSA. Participants at the meeting included four national experts in cost and outcome analysis with backgrounds in mental health and substance abuse, as well as several RAND staff members with experience in cost and outcome analysis. Also participating were staff from SAMHSA, the Casey Family Programs, the SESS Data Coordinating Center, and two of the SESS program sites. The proceedings from the meeting are summarized in the following document:

> Cannon, Jill S., Lynn A. Karoly, and M. Rebecca Kilburn, *Directions for Cost and Outcome Analysis of* Starting Early Starting Smart*: Summary of a Cost Expert Meeting*, CF-161-TCFP, Santa Monica, California: RAND, 2001.

This research is funded by the Casey Family Programs. The opinions expressed and conclusions drawn in this report are the responsibility of the authors and do not represent the official views of the Casey Family Programs, SAMHSA, other agencies, or RAND.

CONTENTS

Preface iii
Figures ix
Tables xi
Summary xiii
Acknowledgments xxxi
Acronyms xxxiii

Chapter One

INTRODUCTION 1

Chapter Two

OVERVIEW OF COST AND OUTCOME ANALYSIS 5
  General Framework 6
    The Policy Analysis Scorecard 6
  Selecting Policies and Impacts 9
    Broad Versus Narrow Formulation of the Problem 10
    The Baseline and Alternative Policies 10
    Considerations in Selecting Impacts 11
  Filling the Cells in the Scorecard 13
    Express Impacts in "Natural" Units 13
    Record Cost Elements as Resource Quantities, Not Dollars 13
    Many Entries May Have to Be Calculated 14
    Explicitly Address Statistical Uncertainty 15
    Explicitly Address Scenario Uncertainty 17

Account for Time Path of Benefits and Costs by Discounting 17
Comparing Policies 18
Common Methods for Aggregating Impacts 19
Aggregating Impacts Has Disadvantages 24

Chapter Three
ISSUES IN COST AND OUTCOME ANALYSIS OF EARLY CHILDHOOD INTERVENTION PROGRAMS 27
Measuring Costs of Early Childhood Intervention Programs 27
Some General Principles of Cost Measurement 28
Types of Costs 29
Capturing the Sources of Cost Variation 32
A Brief Hypothetical Example of Cost Elements and Data 34
Outcome Domains and Measuring Benefits 36
Measuring the Impact of Early Childhood Intervention Programs 37
Translating Program Impacts into Dollar Benefits 43
Comparing Costs and Benefits 45

Chapter Four
BENEFIT-COST FINDINGS FOR EARLY CHILDHOOD INTERVENTION PROGRAMS 49
The Perry Preschool Program 50
Program Benefits 50
Cost-Benefit Analysis 54
The Elmira Prenatal/Early Infancy Project (PEIP) 57
Program Benefits 58
Cost-Benefit Analysis 59
The Chicago Child Parent Centers 63
Program Benefits 64
Cost-Benefit Analysis 65
Lessons for Future Cost-Benefit Analyses of Early Childhood Programs 67

Chapter Five
APPLYING COST AND OUTCOME ANALYSIS TO THE *STARTING EARLY STARTING SMART* PROGRAM 73
The SESS Program and Evaluation Design 73
Using the Scorecard as a Framework 76

    Defining the Baseline and Alternative Policies 77
    Describing SESS Sites 82
  Collecting and Analyzing SESS Program Costs 84
  Collecting and Analyzing SESS Program Benefits 88
  Comparing Costs and Benefits of SESS 95
    Cost-Benefit or Cost-Savings Analysis 96
    Cost-Effectiveness Analysis 97
    Replication Analysis 97

Chapter Six
  CONCLUSIONS 99

Appendix

A. *STARTING EARLY STARTING SMART* GRANT SITES 103
B. SESS PROGRAM ACKNOWLEDGMENTS 127
C. MISSION STATEMENTS OF THE NATIONAL
   COLLABORATORS 129

Bibliography 131

FIGURES

S.1 Some Early Childhood Interventions Have Been Shown to Have High Benefit-Cost Ratios xxiii

TABLES

S.1 Early Childhood Intervention Program Benefit Domains and Illustrative Measures xx
2.1 Illustrative Scorecard 7
3.1 Hypothetical Examples of Cost Elements for Baseline Program and New Program 35
3.2 Early Childhood Intervention Program Outcome Domains and Illustrative Measures 39
4.1 Measured Outcomes and Results for Selected Early Intervention Programs 51
4.2 Costs and Benefits: Perry Preschool Program 57
4.3 Costs and Benefits: Elmira PEIP 62
4.4 Costs and Benefits: Chicago Child Parent Centers . . . 67
4.5 NPV of Benefits Minus Costs for Selected Early Childhood Intervention Programs 68
5.1 Characteristics of SESS PC and EC Demonstration Sites 74
5.2 Illustrative Scorecard for Hypothetical SESS Example: Alternative Policies 79
5.3 Illustrative Scorecard for Hypothetical SESS Example: Program Descriptors 83
5.4 Illustrative Scorecard for Hypothetical SESS Example: Cost Elements 87
5.5 Illustrative Scorecard for Hypothetical SESS Example: Outcomes 92
A.1 SESS Grant Sites 104

SUMMARY

Agency and program administrators and decisionmakers responsible for implementing early childhood intervention programs are becoming more interested in quantifying the costs and benefits of such programs. Part of the reason for this is that foundations and other funders are putting more emphasis on results-based accountability. At the same time, arguments for the value of early childhood intervention are being made within the public sphere on the basis of published estimates of costs and benefits. Program implementers are naturally attracted by statements that a certain intervention produces $4 in savings for every $1 it costs and would like to make similar statements about their own programs. Meanwhile, decisionmakers without particular interest in any given program would like more quantitative decision aids when it comes time to choose among a variety of possible program models or program improvements to implement.

Our objective here is to offer assistance to decisionmakers and program implementers considering an assessment of costs and outcomes. We do not offer a specific step-by-step manual, but we discuss the kinds of issues that must be taken into account and why. We do so in enough detail that readers can decide if this type of quantitative analysis is the right course for them and, if so, can knowledgeably interact with an expert cost-outcome analyst. While we understand that some readers will want to undertake analysis of costs and outcomes to justify a program in which they have a special interest, we take the viewpoint here of an unbiased allocator of funds. What evidence should such a person want to see before concluding that a particular intervention is a wise investment? That sort of evidence is what the implementer seeking to justify further funding will need to present.

We begin by setting the conceptual framework within which program costs and outcomes may be understood. We then draw out some of the implications of that general framework for the analysis of early childhood interventions in particular. After reviewing some examples of such analyses, we apply the methodology to an actual case in which a consortium of program funders must decide whether to proceed with an assessment and, if so, what kind of assessment to undertake. The consortium is led by the U.S. Substance Abuse and Mental Health Services Administration and the Casey Family Programs, and the intervention of interest is the *Starting Early Starting Smart* program.

THE COST AND OUTCOME ANALYSIS FRAMEWORK

Decisionmakers and program implementers just beginning to think about analyzing costs and benefits are often surprised to learn that several analytic avenues are open to them. Which one or ones they choose will have important implications for what they learn and how much they must spend to learn it. Among the choices are these:[1]

- **Cost-benefit analysis** (or benefit-cost analysis) entails comparing a program's benefits to a stakeholder with its costs to that stakeholder. Such a comparison requires putting benefits and costs in comparable terms, and the terms conventionally chosen are dollars. Benefits that cannot be expressed in dollar terms cannot be compared in this manner and are included only in associated qualitative discussion. Cost-benefit analysis seeks to help in deciding whether a program is of value to the stakeholder. Often cost-benefit analysis is conducted from the perspective of society at large.[2]

---

[1] Terminology in this field has not been standardized, and these terms appear in the literature with a variety of different meanings. We have chosen typical definitions.

[2] Of the four analytic approaches listed here, cost-benefit analysis is subject to the greatest challenges in execution and interpretation. That is because benefits must be denominated in dollars, and that adds another source of uncertainty and potential

- **Cost-savings analysis** is restricted to the costs and benefits realized by the government as a whole or a particular funding agency. Only the costs to the government are taken into account, and the benefits are those expressible as dollar savings somewhere in the government. This kind of analysis is used to determine whether a publicly provided program "pays for itself" and is thus justified not only by whatever human services it may render but also on financial terms alone.
- **Cost-effectiveness analysis** determines how much must be spent on a program to produce a particular outcome (or, what is equivalent, how much of a particular type of benefit will result from a given expenditure). While this can be done for multiple outcomes, no attempt is made to sum the complete array of benefits into a single aggregate measure.
- **Cost analysis** alone (no measurement of benefits) can be useful to decisionmakers for a variety of purposes, for example, discovering which factors need to be considered in replicating a program elsewhere or for informing budget projections.

In deciding which avenues to pursue, the decisionmaker or implementer must choose what he or she wishes to learn and consider the funds available for undertaking the analysis. The analyses above are ordered in terms of how much attention must be paid to quantifying outcomes and expressing them in dollar terms (from a lot at the top to none at the bottom). Other variables being equal, the resources and calendar time devoted to the analysis will drop with each successive approach down the list.

As we describe them here, these cost and outcome analysis methods are used only as components within a broader decision support framework that we call *policy analysis* or *policy scorecard analysis* (the latter term derives from the use of a tool called the *scorecard*).[3]

---

disagreement over quantities. For some benefits, dollar conversions are not really feasible. Cost-benefit assessments can thus rarely be comprehensive.

[3] The term *policy analysis* was originally adopted by RAND analysts and others to describe an approach for quantitatively analyzing management problems. Today, the term is used even more broadly to characterize a wide range of quantitative and qualitative approaches to addressing policy issues. Hence, we will employ the more focused term *policy scorecard analysis* for the remainder of this summary.

Despite the name, it does not pertain only to high-level public policies but also to decisions made regarding specific strategies and programs. Policy scorecard analysis offers a framework within which to consider multiple benefits, as required in the first two approaches listed above, and multiple costs, as required by all four. Policy scorecard analysis also entails consideration of alternative programs. This is important for benefit and cost analysis. In trying to determine whether the numbers emanating from these analyses support (further) investment in the program, funders will be asking, "Compared with investment in what else?" A benefit-cost ratio of 1.5 to one ($1.50 of benefits for every dollar of costs) may not be good enough if an alternative with similar objectives has a ratio of two to one. Decisionmakers will thus be considering a range of alternative interventions or at least a choice between funding the program in question and some default course of action (which could be leaving things as they are).

The results of a policy scorecard analysis can be summarized in a simple tool called a *scorecard*. The scorecard lists benefit and cost categories down the side, together with the program design features influencing them, and lists the alternative courses of action across the top. Thus, each cell in the scorecard gives a particular cost or benefit (or design feature) for a particular program. In identifying the row and column heads and filling in the cells—that is, in conducting the policy scorecard analysis—several guidelines must be kept in mind:

- **Designate which benefits and costs accrue to which stakeholders**. If you say that a program generates more savings than costs, people will want to know, savings to whom? And costs to whom?
- **Define explicitly the period over which the analysis applies**. If the purpose of the analysis is to determine whether a program has a favorable benefit-cost ratio or pays for itself in government savings, it is better to look well into the future. No one period or duration is correct, however. The choice depends on the patience of the decisionmaker in question, with individuals typically having shorter planning horizons than society as a whole. This distinction makes a difference because the costs of early intervention programs typically accrue over a matter of months or a few years, whereas the benefits are often not fully realized

until the participating children age into adulthood. Counting such benefits directly entails long-term follow-up of program subjects, though some future benefits can be predicted on the basis of shorter-term trends.

- **Discount future costs and benefits.** Although it is important to count future benefits (and costs), they cannot be counted at full, nominal value. People discount future benefits and costs: getting a $1,000 benefit five years in the future does not look as attractive as getting it now; having to pay $1,000 five years in the future does not seem as onerous as having to pay it now. A real annual discount rate of 3 percent to 6 percent is typically applied to future benefits and costs.

- **Record cost elements as resource quantities.** Until the figures are added up at the end, costs should be recorded in terms of resource quantities—hours of labor, square footage of rental space, etc.—rather than in dollar terms. Prices for these resources can vary from one site to another, and on-budget dollars in particular do not always reflect total costs. A physician may donate time on the weekends, but from society's point of view, that time is not "free"; perhaps it could have been put to another, more beneficial use.

- **Address uncertainty.** Future benefits and costs cannot often be predicted with great confidence. Where a range of values is plausible, that range should be made explicit in the analysis. Likewise, structural uncertainty (e.g., about possible future changes in laws relevant to a program) should also be considered.

The final step in the cost and outcome analysis is to add up all the benefits (or savings) and add up all the costs and compare them across programs. The four analysis methods listed above offer alternative ways of performing this step. Cost-benefit and cost-savings analysis each provide a single measure of merit for each alternative; the alternative with the greatest merit according to this measure is declared the winner. Cost-effectiveness analysis provides multiple measures of merit. They can be combined into a single measure (e.g., the ratio of effectiveness to cost, if a single effectiveness measure dominates), which will be used in the same way as a cost-benefit or cost-savings measure. Or they can be used to define a different kind of selection rule, one that deems "best" the policy that achieves a specified level of effectiveness at lowest cost (a *constant effectiveness* analysis) or that achieves the greatest effectiveness for a given cost (a *constant cost* analysis).[4]

Comparing costs and benefits may not produce a single "answer" that one program is obviously preferable to another. One program may produce a net benefit to one group of stakeholders, while another benefits a second group. The net benefit of one program may be somewhat higher than that for another, but the uncertainty ranges may overlap so much that the advantage cannot be asserted with confidence. Some change in the institutional environment, e.g., tax reform, could shift benefits and costs enough to change the advantage from one program to another. Such possibilities do not subtract from the value of the cost and outcome analysis. On the contrary, some of the most valuable insights are suggestions for policy changes that reallocate benefits across stakeholder groups so that all of them gain and thus have no incentive to block a program.

In most studies, the majority of the analytical effort will come from learning about the domain, structuring the models of how the intervention works, collecting and cleaning data, etc. In short, filling in the scorecard is challenging. Given that groundwork, computing the summary evaluation metrics is straightforward, whether that metric is a benefit-cost or a cost-effectiveness ratio.

Hence, instead of suggesting that one must choose to implement one of these four analysis approaches, it is more accurate to say that one must choose whether or not to conduct a careful, quantitative summation of the effects of the program. If the answer is yes, then there follows a choice of whether to present the results of that analysis to decisionmakers, as a benefit-cost ratio, cost-effectiveness ratio, cost-savings ratio, cost-only analysis, or some combination thereof.

It is thus important to keep cost-benefit analysis, cost-savings analysis, and other forms of cost and outcome analysis in their place. In any decision, some factors can be resolved only through a decision-maker's values and subjective judgment or through negotiation among stakeholders. Likewise, the public quantifying of decision factors may occasionally be problematic (e.g., when an auto manufacturer compares the cost of a safety improvement with the dollar-equivalent benefit of the lives that could be saved by that design change). Nevertheless, these methods can provide valuable input to choosing among different programs, demonstrating a program's worth, improving programs, and replicating them.

---

[4] The latter is sometimes called a *constant budget analysis*, but this is only appropriate if all the costs appear in the budget of the agency making the decision. In many programs, costs may be distributed across many stakeholders. They will not all appear in any single party's budget.

APPLYING THE FRAMEWORK TO EARLY CHILDHOOD INTERVENTIONS

Early intervention programs attempt to improve child health and development by providing young children and their families various social services and supports. Such programs can have effects in four domains: emotional and cognitive development, education, economic well-being (in terms of public assistance, income, and crime), and health. Specific examples of possible benefits within each of these categories are given in Table S.1. Which benefits are measured depends on the purpose of the analysis. Cost-benefit and cost-savings analyses typically seek a comprehensive accounting of the benefits to society or to government (respectively), although many benefits are difficult to express in dollar terms and therefore cannot be aggregated in the cost-benefit assessment. While cost-effectiveness analysis can in principle be performed for any outcome, it is often the case in practice that a single benefit or a narrow set receives most of the attention. A full analysis of the benefits of an early intervention program should include collection of data on as many potential benefits as the analyst's resources permit.

Note that early childhood interventions can benefit parents and other caregivers while simultaneously helping children. It is important to measure benefits to caregivers, because these are often realized over much shorter time periods than are those accruing to children. Ignoring these benefits means underestimating a program's benefit-cost ratio or its potential net savings to government, particularly over the short term—and for some analyses, it will only be feasible to make short-term measurements.

Table S.1

Early Childhood Intervention Program Benefit Domains and Illustrative Measures

	Illustrative Measures for:	
Benefit Domain	Child	Parent/Caregiver
Emotional and cognitive development	Socioemotional and behavior scores IQ test scores Teacher's ratings	Quality of parent-child relationship Quality of home environment
Education	Achievement test scores Grades *Grade progression (repetition)* *Participation in special education* *Educational attainment*	*Educational attainment*
Public assistance receipt, income, crime	*Receipt of public assistance* *Employment* *Earnings/income* *Criminal activity* *Contact with criminal justice system*	*Receipt of public assistance* *Employment* *Earnings/income* *Criminal activity* *Contact with criminal justice system*
Health	Physical and mental health status *Child abuse and neglect* Substance abuse Fertility control *Emergency room visits* *Other health care use*	Physical and mental health status Family violence Substance abuse Fertility control

NOTE: *Italics* indicate measures more easily expressed in dollar terms.

Any analysis of benefits of a program under way *must* include a comparison group. This is a group of children and caregivers not enrolled in the program but similar in as many ways as possible to the program participants and whose progress along the various benefit measures is tracked.[5] Children in particular have a tendency to improve along various measures of development as they grow. Evaluators must take care to ensure that the program benefits they measure are net of what would have occurred naturally or what children would realize anyway from outside influences without the program. Measurements of the comparison group provide estimates of benefits that would have accrued in the program's absence.

---

[5] Ideally, one should randomly assign children and caregivers to program participation versus the comparison group. This ensures that the participation and comparison

Data on progress along benefit measures can be collected by survey questionnaires, tests, or other means of direct interaction with the children and their caregivers. For some benefit types (e.g., reductions in involvement with the criminal justice system), administrative data may be available. When only a few years of data collection are feasible, a glimpse into the future can be obtained through mathematical models that can predict future criminal activity or future earnings on the basis of childhood information. (This cannot of course be done with confidence for any given child, but results obtained for a group of children may be sufficiently reliable for the purpose.)

As with benefits, the cost elements to be included in an analysis depend on its purpose. For example, costs that accrue to society but not to a funding agency are included in a societal cost-benefit analysis but omitted from a cost-savings analysis. Regardless of the analysis to be performed, program costs must be estimated as net of those accrued by comparison group children for similar services. For example, if an intervention is intended to increase prenatal care, the analysis should include only the resources devoted to the visits and services received by program participants in excess of what they would have received anyway (i.e., in excess of those received by the comparison group).

Estimation of costs should follow the general guideline given above regarding the need to estimate resource quantities instead of dollars and to account for "opportunity" costs and other off-budget resource expenditures. Costs borne by participants should also be included, as well as costs borne by other agencies or service providers. Collect-

---

groups are (statistically) identical in both measured and unmeasured characteristics. When the comparison group is selected by random assignment, it is often called a control group. When random assignment is not feasible or desirable, a comparison group can still be chosen, by identifying children and caregivers who are similar in various measured ways to the program participants.

ing cost data for the same set of service providers for both the treatment and control groups allows the analyst to detect both cost shifting (e.g., from one payor to another) and cost offsets (e.g., reduced utilization of services in one area as a result of increased service use in another). In implementing a program, it may also be useful to distinguish between the fixed costs that are not dependent on the number of children served and the variable costs that are. The split between fixed and variable costs will influence the calculation of benefit-cost ratios, net savings, and cost-effectiveness ratios for programs when scaled up to serve larger numbers of children.

SOME ILLUSTRATIVE ANALYSES

Given the challenges and requirements outlined so far, it should not be surprising that not many scientifically sound cost-benefit and cost-savings analyses of early childhood intervention programs with long-term follow-ups have been conducted. Among those recently analyzed or reanalyzed are the following:

- The **Perry Preschool** program provided center-based classes and teacher home visits for one or two school years to 58 children ages three or four in Ypsilanti, Michigan, from 1962 to 1967. Benefits were tracked for both the participants and the comparison group (65 children) through age 27. Benefits included better school performance, higher employment, less welfare dependence, and lower involvement in criminal activity on the part of participants. The most recent cost-benefit assessment evaluates benefits expressible in monetary terms at $50,000 per child, half of that in the form of savings to government, versus a program cost of $12,000 per child (see Figure S.1).
- In the **Prenatal/Early Infancy Project** (PEIP) in Elmira, New York, nurses started visiting mothers when they were pregnant and continued until their child was age two. The objective was to improve pregnancy outcomes and parenting skills and link the mother with social services. Between 1978 and 1980 the program reached 116 first-time mothers. They and another 184 in the control group have been followed through age 15 of the firstborn child. Benefits for the mothers included better pregnancy behaviors and less child abuse in the short term and lower wel-

RANDMR1336-S.1

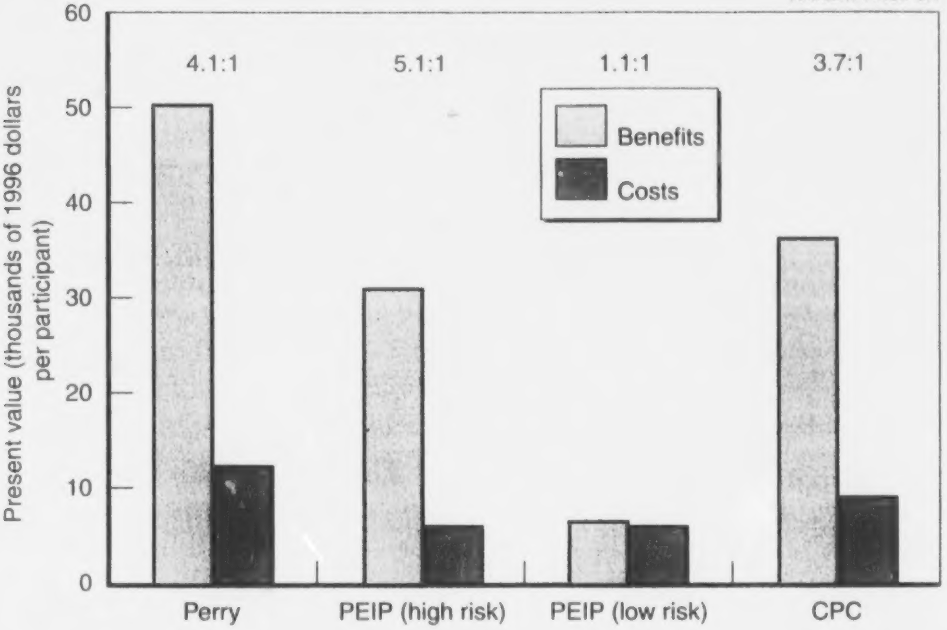

Figure S.1—Some Early Childhood Interventions Have Been Shown to Have High Benefit-Cost Ratios

fare participation and criminal behavior in the long term. The children benefited as well in several domains. For the higher-risk portion of the sample (unmarried mothers with low socioeconomic status), benefits amounted to almost $31,000 per mother-child pair, with almost half of that in the form of a reduction in welfare received by the mother. For the lower-risk portion of the sample, however, benefits came to only $6,700. Program costs were about $6,100.

- The **Chicago Child-Parent Centers** have promoted reading and language skills, provided health and social services, and promoted parent involvement for children in preschool through third grade. A cohort of 989 children completing kindergarten in 1986 was tracked to age 20 and compared with a no-preschool group of 550 children. The program resulted in long-lasting educational-achievement benefits. Higher between-grade promotion rates, reduced special-education use, increased earnings expected as a result of better educational performance, and lower involvement with the juvenile justice system translated into about $35,000 in benefits per program participant. The program cost nearly $10,000 per participant.

These analyses demonstrate that early childhood interventions can generate savings to government and benefits to society that exceed program costs. Indeed, for most of the samples reported above, benefits were a multiple of costs, and all of these programs resulted in benefits that could not be translated into costs and were thus omitted. Therefore, decisionmakers and implementers thinking about performing analyses of costs and benefits should not give up merely because they don't see how some of a program's principal benefits can be converted to dollar terms.

Two further lessons for cost-benefit analysis may be drawn from these examples. First, many important benefits can only be captured through an extended time horizon. The savings from Perry Preschool, for example, did not accumulate to match the level of program costs until the participants were 20 years old. Some of these benefits can be predicted on the basis of shorter trends, but not all can, and confidence in predicted results increases as follow-up periods lengthen.

Second, programs can be beneficial to caregivers as well as to children. In fact, when time is lacking for lengthy follow-ups or when they are not feasible, measuring benefits to caregivers can result in early favorable benefit-cost ratios and net savings. The Elmira program was the only one of those summarized that measured caregiver benefits, and, in that case, savings sufficient to balance costs were tallied within two years of the end of program services.

FRAMING A POLICY SCORECARD ANALYSIS FOR A SPECIFIC PROGRAM

The *Starting Early Starting Smart* (SESS) program is intended to test the effectiveness of integrating mental health services and substance abuse prevention and treatment into early childhood education or primary health care for children from birth to age seven. The program is under way at 12 sites nationwide, seven using the early childhood (EC) education model and five using the primary care (PC) paradigm. (See the appendix for a description of each state.) Most of the sites serve between 100 and 300 children, and comparison groups average out to similar numbers.

By "effectiveness," the program means increased access to, use of, and satisfaction with behavioral health services and increased social, emotional, and cognitive functioning on the part of served children. Data on these benefit measures are being collected over an 18-month follow-up period at intervals that average six months (PC sites) or nine months (EC sites). No cost data are being gathered in this first phase of the program, but a second phase is being planned, and part of that planning is to assess the feasibility of cost and outcome analysis.

SESS program implementers are wise to take cost and benefit evaluation issues into account in the planning stage. Too often, evaluation is considered only after program design has been finalized along lines that preclude sound cost and benefit assessment. SESS's Phase I design raises issues that need to be resolved for Phase II if cost and outcome analyses are to be possible. One issue, for example, is that some sites did not use random assignment (primarily EC sites), which raises concerns about the validity of the treatment group versus comparison group difference as a measure of the true effects of the program. Future demonstration sites should aim for random assignment if at all possible. Another concern is that a few sites are experiencing relatively high dropout rates, which could bias benefit estimates if those who are lost to follow-up are different from those who remain in the study and if they differ in important ways that cannot be observed. Obtaining a consistently high follow-up rate across sites would need to be a priority in Phase II. Also, Phase I has been characterized by between-site variations in services. This is problematic from an evaluation standpoint for a couple of reasons: It complicates interpretation of results, and it complicates the design of comparison groups.

The design of comparison groups for SESS offers lessons for other programs. Because SESS attempts to integrate behavioral health services into existing early childhood and primary care settings, only the benefits of the new, integrated services plus increases in the "dosages" of existing services may be credited to SESS, not the full benefits realized from participation in the early childhood program and primary care. Similarly, only the costs associated with these incremental activities should be considered. Therefore, the comparison groups must be designed to isolate the SESS effects by including everything except SESS. The appropriate comparison groups for this evaluation would consist of children involved in early childhood and primary care programs without the integrated SESS services, not children receiving no services at all.

In the policy analysis scorecard, then, the columns would correspond to the early childhood program without SESS, primary care program without SESS, and then the integrated EC plus SESS and PC plus SESS interventions, along with whatever variants are retained. The rows would be the program descriptors and cost and benefit categories. The program features reported would be those having implications for costs or benefits, e.g., population served, eligibility criteria, age of children at enrollment, qualifications of program personnel, types and "dosages" of services rendered, transportation provisions, and so on. In future demonstrations, this information can be collected through site visits and other mechanisms currently being used in the evaluation of Phase I.

Cost estimates would begin with the cost of serving one child (or child's caregiver) in terms of labor hours expended with the child and in preparing for the session and in terms of materials consumed. These would then be multiplied by dosage per child and number of children served. Fixed costs unrelated to number of children served, such as space rental, would then be identified. Multiplication by unit costs to convert to dollars would be done last. Ultimately, the cost information should be as comprehensive as possible and comparable across demonstration sites.

Benefit measures now being collected for SESS include information on child problem behavior and social skills, child cognitive development, parent-child interaction, caregiver stress and negative or positive behaviors, caregiver mental health problems, caregiver education and employment, and home environment. As discussed above, the emphasis on both child and caregiver benefits will be important to making the short-run benefit tally as complete as possible. Almost all of these measures, however, are within the domain of emotional and cognitive development and are not easily expressed in dollar terms. This makes a formal cost-benefit or cost-savings analysis problematic in that only a limited set of outcomes might possibly be valued in dollar terms to be compared with program costs. Unless the program impact for those outcomes valued in dollar terms is very large and favorable, so that sizable dollar benefits are generated, a cost-benefit analysis would be unlikely to show a favorable outcome for the SESS program based on the information available after two years.

While not the program's main intent, other benefits could result from it. Some of these benefits, in such areas as physical health, labor market outcomes, and involvement with the criminal justice system, could be more easily expressed in dollar terms than those now being measured. These outcomes could be collected for parents or caregivers in the short term, and with longer-term follow-up, for the participating children. If behavioral changes are large in these areas as a result of the SESS intervention, they can produce sizable dollar benefits that, even when discounted, will be a large offset to the costs of the program. This is especially relevant for changes in parental behavior that can be measured even in the short run. Improvements of adult economic and health outcomes have been demonstrated to produce substantial short-run benefits in other early childhood programs.

Costs and outcomes would be measured for both the participant and comparison groups, with the difference between the two constituting the incremental cost and benefits from implementing SESS. To compare the present values of all costs and benefits, it will be important to predict how they will accrue over time. Costs and benefits should also be categorized according to which groups incur them. It will be of interest, for example, to know how much the intervention costs and benefits participants, the agency implementing the program, other agencies, and society as a whole.

Taking all these steps would be sufficient to support as full a cost-benefit or cost-savings analysis as is likely to be feasible given the current state of the art. If SESS decisionmakers wish to be able to say something about the value the program returns to society relative to its costs, the preceding array of evaluation tasks and program design modifications would be required. If they decide it is enough to be able to say how much the program saves the government relative to what it costs, then some elements—costs to participants or losses to crime victims, for example—can be omitted. The overall level of effort required, however, is not likely to change very much.

If SESS funders or implementers would like instead to focus on one or a few prominent measures of effectiveness to compare the different SESS variants with each other, a cost-effectiveness analysis should be sufficient. By collecting cost data, along with data on that one or those few benefits, it would be possible to say, for example, how much child problem behavior decreased (relative to no SESS) per thousand dollars spent on SESS plus EC or SESS plus PC. No conversion of the benefit to dollar terms would be necessary.

Finally, if the purpose was to find out how much program modifications or proliferation of sites would cost, no benefit data would be necessary at all. Clearly, program decisionmakers may have to make trade-offs between what they might like to achieve and how much of a resource commitment they are willing or able to make.

CONCLUSIONS

The recommendations we offer specific to the SESS program may be framed as a set of more-general guidelines for decisionmakers considering cost and outcome analysis of an early childhood intervention program. In particular, among the recommendations that can be applied more broadly are the following:

- Regarding the design of a program evaluation and cost and outcome analysis:
    - Specify the explicit goals of the cost and outcome analysis to guide the scope of cost and benefit data collection and analysis.
    - Identify comparison groups and track the same cost and outcome measures for both comparison and participant groups. If possible, use random assignment to define comparison groups to provide a more valid test of intervention program effects.
    - To minimize attrition in a longitudinal study, devote resources to retaining study subjects.

  - Collect information on program features through site visits and other mechanisms to accurately characterize features of the intervention models as they are implemented and to ensure fidelity to the program model.

- Regarding the collection and analysis of cost data:
  - Collect cost information for both treatment and comparison groups at each site where the intervention program is implemented.
  - Ensure that the cost information is as comprehensive as possible: Costs borne by various parties should be differentiated, the period in which costs are incurred should be identified, and direct and indirect costs, fixed and variable costs, and goods and services provided in-kind should be measured.
  - Plan for proper training and technical support of implementation sites and any cross-site data collection organizations to ensure uniformity in the collection of cost data. Collect information on the cost of data collection, training and support, and the related analyses of the data.

- Regarding the collection and analysis of outcome data:
  - If cost-benefit or cost-savings analysis is the goal, include in the outcome data information for parents and other caregivers in the short term and long term and for children in the long term in those domains with outcomes that can be readily evaluated in terms of dollars and can produce large dollar benefits. The choice of specific outcome measures should be guided by findings from related evaluation studies whenever possible.
  - Obtain information from participants that facilitates collection of administrative data and allows effective tracking of individuals to increase response rates at later follow-ups.
  - Where possible, collect complete histories using retrospective survey questions or administrative data for outcomes that may generate a continuous flow of dollar benefits (e.g.,

labor market outcomes, social welfare program use, use of costly health or education services).

- When supported by other empirical evidence, project future benefits based on observed outcomes. Consider additional method development that would permit such forecasts for a broader range of outcomes.

While we believe these principles are quite general, ultimately these recommendations should be viewed as guidelines that may need to be tailored to the specific circumstances of a given intervention program and its evaluation design. In the end, the objectives of a program's decisionmakers will dictate the shape of the analysis.

The general policy scorecard analysis tools considered in this report, and those specific to cost and outcome analysis, have great promise for improving decisionmaking with respect to such investment programs as the early childhood interventions represented by SESS and its counterparts. When used with skill and judgment, the application of these methods to other programs, such as SESS, will further broaden our base of knowledge regarding the value of these investments and aid decisionmakers in their choice among program alternatives.

ACKNOWLEDGMENTS

We thank our sponsors at the Casey Family Programs and the Substance Abuse and Mental Health Services Administration (SAMHSA) for their support of this project. We especially acknowledge the valuable guidance we received from Peter Pecora at the Casey Family Programs, and Michelle Basen and Patricia Salomon from the Office on Early Childhood at SAMHSA. We also benefited from discussions with a number of cost and outcome analysis experts who specialize in mental health and substance abuse interventions. In particular, we thank Anthony Broskowski (President, Pareto Solutions), William A. Hargreaves (University of California, San Francisco), Brenda Reiss-Brennan (President, Primary Care Family Therapy Clinics, Inc.), and Brian T. Yates (American University). These four individuals served as expert panelists for a two-day meeting convened by RAND in August 2000 to discuss the potential use of cost-benefit analysis or related methods in the analysis of subsequent knowledge development and application studies of the *Starting Early Starting Smart* (SESS) program, a public-private initiative funded and directed through SAMHSA and the Casey Family Programs. This document was shaped by the insights they offered during and after the meeting. In addition, we benefited greatly from information about the current SESS evaluation provided by Fred Springer and his colleagues at EMT Associates, which serves as the SESS Data Coordinating Center. The perspectives on the SESS program offered by David Deere and Karen Rossmaier from the Russellville, Arkansas, site, and Miriam Escobar and K. Lori Hanson from the Miami, Florida, site were also insightful. Other participants at the meeting from SAMHSA and the Casey Family Programs provided useful input.

We thank RAND colleague James Chiesa for writing the summary of this report and providing us with outstanding comments on an earlier draft. We are also grateful for the careful technical reviews provided by our RAND colleagues Carole Roan Gresenz and Shin-Yi Wu. Other valuable comments on the draft report were provided by Richard Boyle and Mark Friedman. Finally, we also thank Patrice Lester and Claudia Szabo of RAND for their very able assistance with the assembly and production of the document.

ACRONYMS

ACT	Assertive Community Treatment (program)
AFDC	Aid to Families with Dependent Children
CER	Cost estimating relation
CFS	Connect for Success (program)
CPC	Child Parent Centers
CQI	Continuous quality improvement
EC	Early childhood (program)
ER	Emergency room
FSAI	Family Services Agency, Inc.
HOME	Home Observation for Measurement of the Environment (Inventory)
PC	Primary care (program)
PCIT	Parent-Child Interaction Therapy
PEIP	Prenatal/Early Infancy Project
QALY	Quality adjusted life year
SAMHSA	Substance Abuse and Mental Health Services Administration
SES	Socioeconomic status
SESS	*Starting Early Starting Smart* (program)

Chapter One

INTRODUCTION

One of the most pervasive trends in social service delivery at present is the "results-based accountability" movement, whereby service providers are increasingly required to provide concrete evidence that their programs generate the desired outcomes. Providers must justify which programs they implement, which design elements to incorporate into their programs, and who will participate. Social science research provides some information about how these program-design features influence outcomes. Although much remains to be learned, the literature on social services aims to address which interventions and treatments affect outcomes and by how much, which groups of individuals respond best to treatment, and, to a lesser extent, which designs elicit the greatest changes in outcomes.

Cost is another primary driver of decisions regarding program design and implementation. Budgets are limited—how many resources are available to expend on accomplishing the goals? Moreover, rather than simply providing a bound for expenditures, cost considerations influence the entire range of decisionmaking. For example, in deciding which program to implement, a policymaker might choose a program that has three-quarters the success rate of the program with the most successful impact, because the former program costs one-third as much as the latter. Similarly, cost considerations figure prominently into program-design decisions, population targeting strategies, and other fundamental parameters.

Research has offered substantially less guidance on cost-related issues than on outcome-related issues. Evaluations of social service programs rarely include information on the total budget for a particular intervention, let alone details on the cost of various components of the program. Furthermore, policymakers have few opportunities to learn about the typical expenses involved in delivering various types of programs. Not surprisingly, service providers have not received the same scrutiny of their cost performance as they have of their outcome performance. This document takes a step toward filling the gap in information available to decisionmakers about the cost considerations that can inform their decisionmaking. While not as extensive as the outcomes literature, a useful body of research on costs and benefits of programs exists, and we present this information with an eye toward helping policymakers incorporate it into their work. Our objective here is to offer assistance to decisionmakers and program implementers considering an assessment of costs and outcomes. We do not offer a specific, step-by-step manual, but we discuss the kinds of issues that must be taken into account and why. We do so in enough detail that readers can decide if cost and outcome analysis is the right course for them and how to knowledgeably interact with an expert cost-outcome analyst.

In doing so, we focus in particular on the issues as they pertain to a class of social service delivery programs that has received a great deal of attention in recent years: early childhood intervention programs. These programs, while varying widely in their design, typically aim to improve child health and development by providing socioeconomically disadvantaged children and their families with various services and social supports during part or all of the period of early childhood (see Karoly et al., 1998, for a review).

In addition to exploring these issues for early intervention programs more generally, we also demonstrate the application of the concepts to a specific example, the *Starting Early Starting Smart* (SESS) program. SESS is a public-private partnership designed to test the effectiveness of integrating behavioral health services with primary care and early childhood service settings for children from birth to age seven. The program is an initiative of the Office on Early Childhood, Substance Abuse and Mental Health Services Administration (SAMSHA) and the Casey Family Programs, along with several other federal sponsors.

Knowledge about the relationship between costs and outcomes is not only useful for individuals who direct specific programs, but it is also important for developing policy approaches at a more general level. One of the arguments for some types of social services is that they function as an investment: spending money now to prevent poor outcomes reaps returns in the form of reduced expenditures to redress poor outcomes in the future. Obtaining better information about program costs and examining the monetary value of program benefits inform the allocation of resources toward prevention services versus remedial services. Hence, in this report we discuss issues related to valuing the benefits (which may include the avoidance of future costs) produced by intervention programs, in addition to issues related to accounting for program costs.

Early childhood intervention programs are one class of social services that may be particularly amenable to this type of "investment analysis." This is primarily because early childhood is viewed as a critical period for physical, cognitive, social, and behavioral development, and inputs in this period may yield payoffs over the rest of a person's life. In addition to the unique role early childhood plays in the life course, children obviously have more years ahead of them than older members of society. This implies that an intervention in early childhood that can evince sustained positive changes will necessarily reap benefits for a longer period than will treatments given later in the life course.

The next chapter provides a general framework for analysis that addresses both costs and outcomes. It includes a brief primer on various types of cost and outcome analysis: cost-benefit, cost-effectiveness, and related methods. The third chapter discusses issues in cost and outcome analysis specific to early childhood intervention programs, while the fourth chapter reviews the literature on cost and outcome analysis for early childhood intervention programs. Chapter Five applies the concepts described in the earlier chapters to the *Starting Early Starting Smart* program, with specific recommendations regarding the evaluation design and implementation of cost and outcome analysis. The final chapter summarizes the main findings and presents conclusions.

Chapter Two

OVERVIEW OF COST AND OUTCOME ANALYSIS

There is a great deal of enthusiasm for applying "business principles" and "investment analysis" to decisions about funding early childhood interventions. The "discipline" associated with these hard-nosed business management approaches is perceived to be a useful antidote to the often emotional appeals and political rancor that accompany policy discussions and decisionmaking in this area. Irrespective of one's view of the relative merits of such methods as cost-benefit analysis for informing policy, cost and outcome methods have emerged as one of the most prevalent tools in the public policy arena (Adler and Posner, 2000). In fact, many states and the federal government have mandated the use of such methods as cost-benefit analysis as part of the policy calculus for various types of policies (Hahn, 2000).

A variety of terms are used, sometimes imprecisely, to refer to the methods in the general class of cost and outcome analyses, including benefit-cost analysis and cost-effectiveness, among others. This chapter will define and illustrate these various concepts and also point out their limitations.[1] We note at the outset, however, that the art and science of quantitative analysis of management problems is far broader than any one of—or even the entire collection of—these notions.[2]

---

[1] Some useful references for further reading are Gramlich (1981), Keeney and Raiffa (1976), Yates (1996), Mishan (1998), and the June 2000 issue of the *Journal of Legal Studies*.

[2] Other tools include the more mathematically advanced methods of operations research, including Monte Carlo simulation, analysis of risk attitudes, Multi-Attribute

GENERAL FRAMEWORK

Over the years, RAND has developed a structured approach for quantitatively analyzing management problems. Called *policy analysis* or *policy scorecard analysis* and is specifically intended for issues involving complex systems and competing interest groups (stakeholders) with different and frequently conflicting goals (Quade, 1989).[3] Policy scorecard analysis requires one to take a broad, systems view of a problem. The problem formulation must include a wide enough range of impact measures to reflect the concerns and goals of all the stakeholders and a wide enough range of alternative policies to map the major trade-offs among the impact measures. Policy scorecard analysis has been applied to a variety of issues such as water management (Goeller et al., 1977; Goeller and the Pawn Team, 1985; Walker et al., 1993), air quality (Goeller et al., 1973), transportation (Hillestad et al., 1996; Walker et al., 1999), drug policy (Caulkins et al., 1997; Caulkins et al., 1999), education (Benjamin et al., 1993; Park and Lempert, 1998), and early childhood programs (Karoly et al., 1998).

Policy scorecard analysis provides a framework within which one can employ the cost and outcome methods mentioned above. We will begin by describing policy scorecard analysis and then use the framework to distinguish among the various cost and outcome methods.

The Policy Analysis Scorecard

A central construct in policy scorecard analysis is the scorecard (see Table 2.1). This is simply a table with a column for each policy and a row for each impact measure. Where possible, entries in the table should be cardinal measures of the size of an impact (e.g., policy A

Utility Theory, and optimization methods. The methods discussed here are geared toward helping people make choices. Other aspects of quantitative analysis of policies may be more appropriate for other dimensions of management, including program design, budgeting, forecasting, consensus building, marketing, and so on.

[3] The term *policy analysis* was originally adopted by RAND analysts and others to describe a specific systems approach to problem formulation and analysis. Today, the term *policy analysis* is used even more broadly to characterize a wide range of quantitative and qualitative approaches to addressing policy issues. Hence, we will employ the more focused term of *policy scorecard analysis* for the remainder of our report.

costs $125 million per year). But they may be rankings (policy B is first, followed by A, D, and C in that order) or categories (High versus Low or Good versus Intermediate versus Bad) or even text descriptions (policy A has special feature X).[4] To select the preferred policy, the decisionmaker will compare the columns in the scorecard to determine which one he or she prefers. Typically, no policy will beat all the others on every impact measure, so selecting a policy will involve trading off one impact against another.[5]

At the end of a study, a scorecard is often a good way to summarize the results of an analysis to the sponsor. For this purpose, the analyst must restrict the size of the scorecard, so the scorecard will pre-

Table 2.1

Illustrative Scorecard

	Alternative Policies			
Impacts	Baseline: No Program	Program A	Program B	Etc.
Program Descriptors				
Parent training				
Child health screen				
Etc.				
Cost Elements				
Labor paid by agency				
Rent paid by agency				
Participant travel				
Etc.				
Outcomes				
Child outcome A				
Child outcome B				
Parent outcome A				
Etc.				

---

[4] See Caulkins et al. (1999) and Hargreaves, et al. (1998, p. 107), for illustrations of the use of scorecards.

[5] The notion of a trade-off can be illustrated as follows. Anyone would agree that it's better to be rich and healthy than poor and sick! But there may be no way to achieve both objectives simultaneously. One may have to sacrifice some of one to obtain more of the other, for example, by cutting back on work (and hence income) to reduce stress-related disorders.

sent only a handful of alternatives (columns) and impacts (rows) that illustrate the major choices and key trade-offs. At the beginning of a study, constructing a notional scorecard is a useful aid to problem formulation. The major tasks of formulation are specifying the range of alternatives (columns of the scorecard) to be considered, specifying the kinds of impacts (rows of the scorecard) to be estimated, and specifying how those impacts will be measured (entries in the cells of the scorecard). Initial formulation of the problem will typically produce far too many alternatives and impacts to be included in an actual scorecard, and a large part of the analyst's art consists of screening out the less desirable alternatives and the less useful impacts, ending with a scorecard of manageable size that does not mislead the client.

The scorecard is most obviously an appropriate construct for decision problems, such as selecting one program from among several alternatives or designing a program that maximizes the return on investment or that maximizes the effectiveness for a given budget or that minimizes the cost while achieving specified outcomes. Less obviously, the scorecard construct is also appropriate for the task of program evaluation, where at first glance it appears that only one program exists and no alternatives need be considered.

Initial appearances can be deceiving. Most fundamentally, even defining the costs and benefits of a program requires distinguishing what is part of the program from what is not. To say this another way, it requires establishing a baseline, a state of the world without the program that can be compared to the world with the program in place. In clinical trials of a new drug, for example, the baseline is established by a control group of subjects who do not receive the drug. They are compared to subjects similar in all ways except that they are given the drug.

Beyond this, a program is usually evaluated with an eye toward improving it, replicating it in a different setting (e.g., serving a different population), scaling it up, or perhaps canceling it. That is, a program evaluation is generally expected to lead to a decision. A decision to cancel the program will be based on a comparison of the program to the baseline. Improving the program, replicating it, or scaling it up or down will involve comparing the program as currently implemented with one or more variations of the program.

In the remainder of this chapter, we discuss three questions:

- What policies (columns) and impacts (rows) should appear in the scorecard?
- How do we fill out the body of the scorecard?
- Once the scorecard has been constructed and filled out, what methods do we employ to attain our analysis objectives? The methods we will consider are the four listed previously, namely, benefit-cost analysis, cost-savings analysis, cost-effectiveness analysis, and cost analysis alone.

This discussion will proceed linearly, whereas in an actual study the analyst would iterate among these steps. Early in the study an analyst will tentatively select policies to consider but may later discover that information about some policies is simply too difficult to collect. and these policies must be dropped. Or an analyst may discover that none of the policies offer benefits to a particular stakeholder group and try to design a new policy that fills that void. For similar reasons, the analyst may add or delete impacts during the course of the study.

SELECTING POLICIES AND IMPACTS

When someone argues that a program or policy is the "best" way (or even a "good" way) to solve problem *X* (where, for example, *X* is traffic congestion or air pollution or drug abuse or child neglect), an important reaction should be to ask, "Compared to what?" The columns in the scorecard answer this question. Looking at the scorecard, the analyst and the decisionmaker can compare the policies that exist, but they can only speculate about policies that have been omitted.

A second important question is, "How do you measure the 'goodness' of the policy?" Or to say this another way, what are the costs, the products, the side effects, the unintended consequences? The rows in the scorecard answer this question. The analyst and the decisionmaker can consider costs or population served or any other impact only if it is included in the scorecard.

Selecting the rows and columns of the scorecard is thus a key aspect of a study design, with decisions about whether rows and columns are defined in a more limited fashion or more expansively, along with the specific elements to include in each dimension. We discuss each of these aspects in turn.

Broad Versus Narrow Formulation of the Problem

Formulating a policy problem broadly means including a wide range of alternative policies and impacts.[6] A broad formulation has both advantages and disadvantages. Data gathering and analysis for a wide range of policies and impacts will be more costly, time-consuming, and difficult than for a narrow range. If the choice of a preferred policy is to be made by a group, consensus will be harder to achieve when there are many alternatives to choose from and many impacts on which to compare them. On the other hand, a narrow formulation may exclude impacts that measure important costs and benefits and may ignore policies that excel on the excluded impacts. There was a time, for example, when factories were sited, built, and operated without regard for their environmental impacts.

If the objective of the analysis is to improve an existing policy or to replicate a policy in a new environment, it is important that there be adequate variation among the policies in the scorecard. The role of analysis in this context is to do as much policy improvement or policy adaptation on paper (or by computer) as one can, so that the worst features can be weeded out before the policy is actually delivered to real people.

The Baseline and Alternative Policies

The illustrative scorecard above includes a column for a policy or program labeled "baseline." Typically, this policy represents the world without the alternative policies or programs under consideration. For many program evaluations, the baseline is the control group or comparison group. In experimental evaluations, individuals are randomly assigned to the control group (i.e., the group that receives no new program services or faces the status quo) or the treatment group (i.e., the group that receives the program services or faces the policy alternative). When properly implemented, randomized experimental designs are considered the "gold standard" for evaluation research because the control and treatment groups are as similar as possible except for participation in the program. Thus, any differences in the cells of Table 2.1 can be attributed to the impact of the program or policy. Quasiexperimental designs include a comparison group chosen on the basis of matched characteristics but not random assignment.

---

[6] We distinguish here between a *wide range* versus a *large number* of policies and impacts. It is possible to inflate the number of policies or impacts by including numerous minor variations of either, but this does not increase the breadth of the formulation.

The column corresponding to the baseline also provides a place to record *scenario assumptions*, i.e., assumptions about aspects of the future state of the world that may influence the impacts of the other policies. We will have more to say about scenario assumptions later. The overall objective of the analysis is to compare this baseline to columns representing the various alternative programs or policies and assessing which column represents the optimal choice, given the choice mechanism selected. We will return to the discussion of how to choose among alternatives below.

Typically, all policies save the baseline will be constructed by combining policy elements. For example, a policy element may involve delivering a particular service or intervention (e.g., drug counseling or parenting training) to a specified target population (e.g., low-income first-time mothers in a particular neighborhood) by a certain method (e.g., home visits or sessions at a clinic). Then a policy might deliver different services to several different populations (e.g., parenting training to one group, drug counseling to another). It might deliver different services at different venues. Any particular policy will probably have a fairly well-defined service area, which will be the same for all services it delivers and all populations it serves. Different policies can serve different areas, however.

Considerations in Selecting Impacts

As seen in Table 2.1, the illustrative scorecard includes rows for program design, as well as those capturing cost elements and outcome measures. Both the cost elements and the outcomes should be broken out by stakeholder and by time. Breaking out costs and outcomes by stakeholder means identifying who pays or benefits. This is important because the costs and benefits of a program might accrue to different stakeholders, which is likely to enter the decisionmaking process. For example, a policy or program that benefits group A at the expense of group B will often be opposed by the latter, even if total benefits exceed total costs. Breaking out the costs and outcomes by time means specifying when the cost is incurred or the benefit realized. A policy that incurs costs today but yields benefits only years later may not appeal to a term-limited politician, even though the policy might appeal to somebody with a longer view.

Typically, this implies that the scorecard will have a large number of rows. In many problems it is easy to identify half a dozen stakeholders, e.g., the government agency implementing the policy, two or three other agencies, the target population, family members of the targeted population, and other residents. The analyst will define at least one impact for each stakeholder (e.g., cost) and several outcomes for the targeted population. Each impact may occur this year or in any of the next N years. It can add up to dozens or even hundreds of rows.[7]

Some outcomes may take so long to be realized that they cannot be observed before the decisionmaker must choose a policy. Early childhood interventions are intended, among other things, to reduce the likelihood that the child will drop out of school or use drugs or commit crimes as an adolescent or young adult. A decade or more must pass before we can observe whether these goals have been met. In place of these key but sometimes unobservable impacts, the analyst must substitute short-term outcomes that are reasonable predictors of the more important long-term outcomes. But "reasonable predictors" is a flexible term. It may be that nothing that can be observed within (say) two years has been demonstrated (e.g., by a careful clinical trial) to be a "good" (e.g., acceptable by academic standards) predictor of a future outcome. It is better to include a predictor that is deficient by academic standards than to omit the impact from the analysis. As we discuss later, however, the subsequent analysis must take due account of the impact's uncertainty.

---

[7] The sheer size of the scorecard should not be a cause for dismay. At initial formulation, the scorecard will include many more impacts (and alternatives) than it will toward the end of the study. A major part of the analyst's art is devoted to screening out alternatives and impacts that are not informative. Moreover, for presenting final results to the client, the analyst may split the one scorecard into many, each with a different focus. For example, if the focus is on how the state of the world (e.g., the unemployment rate) affects the performance of different programs, the analyst can construct a handful of scenarios (e.g., "pessimistic," "best guess," and "optimistic") and create one scorecard for each. Or if the focus is on performance in the short run versus the long run, the analyst could construct one scorecard with impacts at one year, another with impacts at five years, and so on.

FILLING THE CELLS IN THE SCORECARD

To complete the scorecard, the individual cells must be completed. In this section, we offer several guidelines to be followed, as well as methodological issues that arise as part of this process.

Express Impacts in "Natural" Units

Entries in the scorecard should be expressed in "natural" units. That is, where possible, they should be cardinal measures of the size of an impact (e.g., policy A costs $125 million per year). However, cardinal measures—those that can be expressed quantitatively in well-defined units—will not always be available. Where necessary, such as for qualitative impacts, entries may be rankings (policy B is first, followed by A, D, and C) or categories (High versus Low or Good versus Intermediate versus Bad) or even descriptions (policy A has feature X). The reason for this advice is that analysis is often criticized for ignoring considerations that cannot be quantified easily (for example, see Sen, 2000). Including difficult-to-quantify impacts (i.e., impacts for which cardinal measures are hard to define) preserves a chance, at least, to include them in the analysis. Even if they can't be included in the analysis except by artificial and labored means, they can nonetheless figure in the deliberations of the decisionmaker.[8]

Record Cost Elements as Resource Quantities, Not Dollars

In particular, cost elements (one of the categories of impacts shown in the illustrative scorecard) should generally be shown as quantities of resources, such as man-years or gallons of gasoline. They should not be expressed directly as dollars unless the resource inventories behind the dollars are unavailable, even though the analyst intends to price them out later during the analysis phase of the study. There are several reasons for this. First, prices differ from place to place. For example, a program implemented in one city may make use of volunteer labor and donated facilities, while a similar program in another city may need to pay for some or all of these resources. Second, resources may be shared, and a reported dollar cost will be based on accounting assumptions about whose budget is charged for how much of the resource. Those accounting assumptions can differ for a program implemented elsewhere. Third, some resources may be hard to get quickly or even hard to get at all. It might be necessary to find an alternative way to do things in order to implement the program in another location. For example, there may be no emergency room available in a rural setting, while there will be one in a city.

---

[8] Analysis has limitations, after all. The analyst does not *replace* the decisionmaker. Rather, he or she collects, processes, and displays information in a way that will *help* the decisionmaker arrive at better decisions.

Many Entries May Have to Be Calculated

Entries in the scorecard can come from a variety of sources. The most obvious is direct measurement, either by the analyst or by others (e.g., an experiment or demonstration reported in the literature). Because few policies in the scorecard will have been implemented in their entirety, direct measurements of their impacts will not exist. Data on the impacts of individual policy elements often will exist, however, and just as a policy is built from policy elements, so too can the impacts of a policy be estimated from the impacts of its elements.

A rather simple model will often serve to estimate the resources employed in a program, as a function of its service area, its capacity (i.e., the number of people the program is designed to serve), and its workload (the number it actually serves). Simple geometric arguments can provide estimates of travel distances, which can easily be converted to travel times (at so many miles per hour) and transportation costs. The workload (number of people served) usually translates easily into direct hours of labor (e.g., so many visits per person served times so many minutes per visit plus travel time).[9]

---

[9] It is important to add in indirect hours as well. For example, in addition to time spent directly delivering a service, a service provider will also spend time completing paperwork or engaged in other administrative tasks required for direct service deliv-

Likewise, model-based estimates of benefits may be possible when direct observation is not available. In some cases, longer-term impacts may be projected based on short-term outcomes using relationships estimated in other studies or derived from meta-analyses. In the early childhood literature, for instance, estimates of adult lifetime earnings have been projected based on observed final educational attainment or labor market outcomes in early adulthood (see Chapter Four). Ideally, these projections reflect the latest understanding in the literature and will acknowledge the degree of sophistication of the models and their acceptance by other analysts.

Indirect costs and benefits—those tangentially associated with the program or services being evaluated—may also need to be estimated. One example of an indirect cost is an increase in the use of pediatric care by a participant in a program that provides other types of early childhood intervention services. To obtain this from actual measurements, one must measure the use of pediatric care by participants and by a control group, and subtract the second from the first. (Data about the control group will help fill the "baseline" column of the scorecard.) In the absence of actual measurements, one might bound the cost by assuming participants will use pediatric care at whatever rate the American Medical Association recommends.

Because ideal data for each entry in the scorecard are not likely to be available, the analyst must use creativity and informed guesswork to fill it in. Rarely will there be enough data of high enough quality that all entries can be estimated with high confidence. Large blocks of entries may need to be based on educated guesswork if they are not to be left entirely blank. Of course this affects the reliability of the analysis, but in our view, it should not be taken as an excuse to abandon analysis altogether (see the discussion in Quade, 1989).

Explicitly Address Statistical Uncertainty

Entries in the scorecard will be uncertain. Some of this uncertainty will be of the familiar statistical variety.[10] Survey results will have an

---

ery. One must be sure to add enough indirect hours. It shocks many people, but it is reasonable to estimate indirect hours as 1.0 to 1.5 times as large as direct hours.

[10] Another source of potential uncertainty is errors in measurement. Data quality concerns are relevant for both cost and outcome measures, and may be an issue with

error of x percent. Estimates from an equation fitted to data by ordinary least squares will have a standard error. The sizes of these errors should be shown in the scorecard so the analyst and decisionmaker can judge whether two policies differ significantly in a particular impact. When available, these errors can also be used with the aggregation methods discussed below to provide estimates of the uncertainty associated in the cost-benefit, cost-effectiveness, and related analysis. In some cases, only a subjective characterization of uncertainty is available, but even subjective characterizations are generally better than providing only point estimates of quantities that are in fact uncertain.

Moreover, one should distinguish between the statistical significance and practical significance of such a difference. If the standard error of an impact's estimate is low, two policies may have a statistically significant but practically inconsequential difference in that impact. By contrast, if the standard error is large, the difference may be statistically insignificant but practically important. In the latter case, it is not known whether the difference is real, but it is important to find out. One case where the error may be large is when a short-term impact has been used as a predictor of an important long-term outcome.[11]

The issue of statistical uncertainty means that sample size considerations are important at the design stage of a program evaluation, both for measuring program impacts and for conducting related cost and outcome analyses. Typically, in experimental and quasiexperimental study designs, sample sizes for treatment and comparison/control groups are chosen by balancing cost and other implementation concerns against the statistical power to detect differences between the two groups. If cost and outcome analysis is planned, the decision about sample size will have implications for the ability to draw inferences about program differences in economic terms as well.

---

information obtained through direct observation, surveys, or administrative sources. Ideally, the most reliable source of data is available for any given scorecard element and any known concerns about data quality are acknowledged by the analyst.

[11] See discussion in Caulkins et al. (1999) for an illustration of how statistical uncertainty can affect and be handled in cost analysis.

Explicitly Address Scenario Uncertainty

In most cases, factors completely outside the policies of interest will affect the sizes of the impacts. For example, five years ago a family enrolled in an early childhood intervention program could, in principle, remain on public assistance indefinitely. Under current law, the family will be dropped from the rolls after a few years. Depending on the scenario, the eventual benefits of helping a mother, or eventually a child, enter the workforce will be quite different. Or suppose the program provides job training (or refers participants for job training). The effectiveness of this service depends on the local availability of jobs, which in turn depends on the state of the economy. Policymakers should be made aware of assumptions about future developments that may drive the success or failure of the program (Dewar, 1993).

Including a baseline in the scorecard provides a vehicle for including scenario assumptions. Frequently an analyst or decisionmaker will talk about the cost or benefit of a policy or program, with no reference to the baseline at all. This is a convenient shorthand, but it suppresses the fact that the costs and benefits depend on more than the features of a policy or program. They depend as well on the environment in which the policy is implemented and the future environment in which it is operated—for example, the population the program is serving or the other services available in an area. Thus, ideally, the analyst describes the baseline in a rich enough manner that it includes all of the assumptions about the future state of the world that are likely to affect the performance of any of the policies. If the analyst anticipates replicating a policy in another environment, the baseline should also include any factor that may differ between the current and target environments, if that factor influences the performance of any policy.

Account for Time Path of Benefits and Costs by Discounting

A final consideration in filling in the cells of the scorecard involves how to value costs or benefits that accrue in the future. For example, suppose that a home visiting program for 100 children would reduce the expected number of emergency room (ER) visits per child in each of the subsequent three years by one visit. If each ER visit costs an average of $200, one might think the benefit is best described as the elimination of one visit per year per child: 100 participants x 1 visit per participant x $200 per visit x 3 years = $60,000. But the usual practice is to weight or value outcomes that occur sooner more than outcomes that are delayed. It is obvious why this should be so with money. One would rather have $1,000 today than $1,000 next year, because if a person had $1,000 today he or she could invest it and have more than $1,000 next year. The same logic of "discounting" or applying "time preferences" can be applied to nonmonetary outcomes, and at the same rate (Keeler and Cretin, 1983).

While there is consensus that future outcomes should be discounted, there is no consensus as to what rate should be used, although 4 percent is typical.[12] If we apply a 4 percent discount rate to this example, we would calculate the "present value" of reducing ER visits as the amount saved per year, scaled by a discount factor, which is 1/(1.04)^N, where ^N indicates that 1.04 is raised to the power based on the number of years in the future the value is measured. In this case, the present value would be: $20,000 + $20,000 x (1/1.04^1) + $20,000 x (1/1.04^2), or $57,700. The term "present value" connotes the idea that given a 4 percent discount rate, one should feel the same about receiving $57,700 today and receiving a savings of $20,000 at the end of each of the next three years. In terms of nonmonetary outcomes, you could discount the 100 ER visits per year for the next three years by the same rate to get a present value of 289 visits. While discounting is a routine method in analysis, to simplify exposition and focus on the more fundamental conceptual issues, it will be suppressed in the remainder of this discussion.

COMPARING POLICIES

Once the analyst has a scorecard with all the cells filled in, it is possible to compare the policies. The purpose of the analysis is likely to be one of the following:

---

[12] In medicine, 3 percent and 5 percent are recommended (Gold et al., 1996). A variety of RAND analyses in the drug, criminal justice, and children and youth intervention policy areas have used a 4 percent discount rate (e.g., Rydell and Everingham, 1994), while Karoly et al. (1998) explicitly consider a range of discount rates from 0 to 8 percent. Rates between 0 and 10 percent or higher have also been used. The choice of rate may be a function of the time preference of the stakeholder or decisionmaker.

- Select the "best" policy (column) in the scorecard.
- Design a new policy that is "better" than any of the policies in the scorecard.

Since policies have many different impacts, it is highly likely that one will be better than its alternatives on some impacts but worse on others. Comparing policies therefore requires trade-offs to be made among the impacts. Analysts often devise metrics that summarize most or all of the impacts into a single, aggregate score. These metrics define trade-offs among the impacts, because a unit improvement in one impact is worth whatever size reduction in a second impact is necessary to keep the score constant.

Not all methods of selecting a "best" policy use a single aggregate measure of merit. One common method, called a *constant-cost analysis*, uses one measure of effectiveness and one of cost and deems the policy "best" that maximizes the effectiveness measure while not exceeding a specified cost. If cost is defined from the point of view of the decisionmaker, it is sometimes called a *constant-budget analysis*. Another method, called a *constant-effectiveness analysis*, permits the use of several measures of effectiveness and one of cost. The policy is deemed "best" that achieves specified levels of each of the effectiveness measures while minimizing cost. These methods are only useful, however, if they rely on a small number of measures. Thus they require the impacts in the scorecard to be substantially aggregated. We now review some of the alternative ways of creating summary metrics of the costs and benefits of policies.[13]

Common Methods for Aggregating Impacts[14]

Cost-benefit analysis converts the benefits and costs into common units, most often dollars, and then notes which is greater. Benefits that cannot be expressed in dollar terms cannot be compared and are excluded from the formal analysis. The purpose of cost-benefit analysis is to help in deciding whether a program is of value to the decisionmaker, or notional decisionmaker, when the analysis is done from the perspective of society at large. The greater the margin by which benefits exceed costs, the better the investment we consider the program to be.[15]

---

[13] We stress that these summary metrics often cannot include all the impacts in the scorecard. Generally they include only quantified impacts (i.e., those with cardinal measures) and sometimes not all of them. Remember, the goal of analysis is to *help* the decisionmaker, not to *replace* him or her.

[14] The terms in this section are common to the field, but considerable variation occurs among commonly used definitions. We have chosen definitions that are typical but not universal.

One distinction among approaches to comparing costs and benefits concerns the stakeholder to whom costs and benefits accrue. *Cost savings analysis* is a term sometimes used to refer to a cost-benefit analysis done from the perspective of the government generally or a particular government agency. It compares only the costs to government and the savings to government generated from a program. Cost savings analysis is used when asking questions, such as whether the benefits of a program to government pay back the costs taxpayers invested in the program.

The two common ways to compare the benefits and costs are by looking at their ratio or their difference. Dividing the benefits by the costs yields a *benefit-cost ratio.* Referring to our example of the home visiting program above, suppose the program cost $300 per child, for a total cost of $30,000. Then, the benefit-cost ratio for the program is $57,700/$30,000 or 1.9. Subtracting costs from benefits yields the net value. Because discounting is often involved, this is most often called the *net present value,* or NPV.[16] In our example, the NPV of the parent-training program is $57,700 – $30,000 = $27,700.

When other program alternatives to this treatment program exist, one should generally choose the program with the greatest measure of merit. For example, if three alternative home visiting programs have NPVs of $15,000, $27,700, and $45,000, respectively, and you can only implement one, choose the last. Note, however, that using the benefit-cost ratio may lead you to choose a different alternative, if the costs of the alternatives are substantially different.

---

[15] It might seem natural to say that if benefits exceed costs, then the program is a good investment. But this ignores the question, "Compared to what?" That is, the question is not whether the investment is "good" in some absolute sense, but whether it is better than the alternatives.

[16] See Karoly et al. (1998) and Currie (forthcoming) for examples of cost analysis of early childhood programs that use the NPV approach.

Cost-effectiveness analysis tries to side-step uncertainties about how to value different aspects of programs by looking at the ratio of benefits to costs without reducing them to common units. For example, our hypothetical home visiting program has a cost-effectiveness ratio of 289 ER visits averted / $30,000 in program costs = 9.6 ER visits averted per thousand dollars spent. The ratio of effectiveness to cost is sometimes informally termed the "bang for the buck." This term comes from cost-effectiveness analysis in the military context, where monetizing outcomes, such as the ability to deliver a given payload of bombs, is similarly difficult. In other contexts, it is common to invert the ratio, calculating the cost per unit of benefit purchased. For instance, health care programs are often evaluated in terms of the cost per quality adjusted life year (QALY) saved (Kamlet, 1992). In those cases, smaller numbers indicate more efficient programs.[17] Whether it is more felicitous to think about maximizing what is obtained for a given cost or minimizing the cost necessary to attain a given effect depends on the context. The term *cost-effectiveness* covers both variants, although calculations of cost per QALY are sometimes called *cost-utility analyses.*

The cost-effectiveness ratio for a single program is often difficult to interpret. Most people do not have an intuitive sense of whether averting 9.6 ER visits per thousand dollars is a lot or a little. But if one calculates the cost-effectiveness ratio for each available intervention, the one with the highest ratio is the preferred place to invest the next dollar. (If the ratios are computed in terms of cost per unit benefit, not benefit per unit cost, then the intervention with the smallest ratio would be preferred.) For example, if alternatives to the home visiting program had cost-effectiveness ratios varying between two and seven ER visits per thousand dollars spent, then the home visiting program would, all other things being equal, seem to be a more appealing place to invest the next thousand dollars.

One can also compare programs in terms of the lengths of time they must remain in operation to recoup the initial investment, sometimes called the *payback period*. Typically, for a given treatment population in the early stages of a program, only costs are generated. Once the program services end for that population, the cumulative costs do not change. During the period of program implementation, benefits may begin to accrue and they can continue to grow after the program services end. For example, Karoly et al. (1998) found that the Elmira home visiting program paid back its costs of delivering services to the treatment group after about two years, while the Perry Preschool Program took nearly two decades to recoup its costs for the cohort it served.

---

[17] See Greenwood et al. (1998) and Caulkins et al. (1999) for examples of cost-effectiveness analysis for early childhood programs.

As discussed above, programs often produce multiple benefits. For example, a substance abuse treatment program might not only reduce cocaine use, but it might also avert a given number of serious crimes and the years of prison time associated with those crimes. Cost-effectiveness ratios per se are limited to a single outcome and so have a hard time fully reflecting such a range of benefits. But, sometimes the candidate interventions produce the various benefits in almost fixed proportions. In that case, focusing on one benefit is not problematic because whichever program generates the most "bang for the buck" with respect to that benefit does so with respect to the other benefits as well.[18] But that is by no means always the case. For example, drug prevention programs reduce the number of cocaine users by a greater proportion than they reduce the quantity consumed; for drug treatment programs, the opposite is true.[19]

When outcomes are produced in different proportions, one may calculate a cost-effectiveness ratio for each important outcome. This is sometimes called *cost-consequences analysis*. For some purposes listing explicitly the set of outcomes produced per thousand or per million dollars invested is useful. For others, decisionmakers may prefer a single, bottom-line summary. Cost-benefit analysis provides that bottom-line summary by reducing all outcomes to a common currency.

---

[18] For example, in Greenwood et al. (1998) incarceration policies tended to produce reductions in different types of crime in constant proportions, so the analysis could usefully focus on one aggregate measure (serious crimes) without worrying about the fact that some types of serious crime (murder) are in some sense more "costly" per offense than are other serious crimes (e.g., robbery).

[19] For other such examples, see Caulkins (2000).

Notice that when some of the benefits are *avoided costs*, as in the example of reduced crime and use of the criminal justice system, ambiguity can arise with respect to the computation of the benefit-cost ratio. The NPV is the same whether the savings in prison costs are counted as a benefit or a cost offset. But the benefit-cost ratio changes. If one counts the taxpayer savings from reduced prison time as a benefit, the benefit-cost ratio will include the prison cost savings in the numerator. If one views it as a cost offset, it is possible that the net cost to the taxpayers of funding the treatment program is zero or even negative (depending on the size of the offset). Thus, it is possible for the benefit-cost ratio to become negative or to be undefined (e.g., when net costs are zero).[20]

That one can compute different benefit-cost ratios depending on whether some outcomes are viewed as benefits or cost offsets leads some observers to recommend focusing on the NPV, not the benefit-cost ratio. However, the NPV may depend on the scale of the project. A mediocre program implemented throughout a large state such as California may have a larger NPV than an outstanding program implemented in a small state. In these contexts, it is thus useful to discuss the NPV per unit of activity, such as the NPV per child or family in a program.

Because monetized physical outcomes are not the same as "real" money, one can make an argument for putting all outcomes that literally involve dollars in the denominator and segregating the "dollar equivalent valuations" in the numerator of the benefit-cost ratio. This approach is sometimes labeled *cost-offset analysis*. When the alternative algorithms suggest different results, the differences should be highlighted and explained.

Finally, *cost analysis* alone, with no accounting for program benefits, can also be useful to decisionmakers for a variety of purposes—for example, discovering which factors need to be considered in replicating a program elsewhere. Compared with a cost-benefit analysis or the related methods that also require measurement and analysis of program benefits, this approach requires the fewest resources to implement, albeit with a corresponding reduction in what is learned about the program's impacts. It is most valuable when it identifies who bears which portion of the costs, not just the total cost.

---

[20] Why would one ever consider avoided costs to be negative costs rather than positive benefits? Because an impact can be negative for policy A and positive for policy B. Whether it is categorized as a cost or a benefit, it will be negative for one policy and positive for the other.

Aggregating Impacts Has Disadvantages

Cost-benefit analysis and the allied methods described above collapse the impacts to a single measure of merit, but policymakers answer to the concerns of particular constituencies—perhaps voters, heads of their agencies, clients of their agencies, and others. Also, for most people, decisions are guided by equity and justice as well as efficiency considerations. In short, distributional issues matter.[21]

If we all agreed how the costs and benefits ought to be distributed among stakeholders, these issues could be incorporated into cost-benefit analyses. One could call improving the lot of criminals a cost rather than a benefit and assign some dollar-equivalent penalty to it. One could decide that from society's perspective, increasing the income of poor people is worth twice as much per dollar as increasing the income of people in the middle class. One could count as an objective not just improving the average lot of people in different neighborhoods, but also reducing the inequity between them (see, for example, Keeney and Raiffa, 1976). But people differ on these matters, so they place different relative values on various outcomes. As a result, different people will rank policies in different orders, and no single measure of merit will satisfy everybody.[22]

For example, it might be less costly to implement a publicly funded daycare program in a middle-class neighborhood than it is in a poor neighborhood, perhaps because it is easier to find buildings that meet asbestos standards or because fewer of the children have special needs. Furthermore, the impact on tax revenues may be more favorable if the middle-class parents who are freed to work would earn more and be taxed at higher marginal rates than the parents in poor neighborhoods would be. Nevertheless, few would openly sanction targeting such government subsidies at privileged rather than at at-risk families.

---

[21] See Posner (2000), Frank (2000), and Richardson (2000), for discussions of distributional issues for cost and outcome analyses.

[22] Of course, if the policy choice were up to a single decisionmaker, he or she would use a measure of merit that reflected his or her views, and a suitably tailored cost-benefit analysis would suffice. Policy choices in the real world are often the product of commitments by a range of individuals and institutions. A famous theorem by Kenneth Arrow (1951) demonstrates, roughly speaking, that there is no analytically defensible way to combine the different preference schemes of multiple individuals to obtain a group preference. Thus, different equally justifiable methods of combining individual preferences can lead to different group preferences. Coming to a consensus, therefore, has to be essentially a political process rather than an analytic one.

A rather infamous example pertains to the value of stolen property. One school (Cook, 1983, and Harwood, Fountain, and Livermore, 1998, are examples) argues that when goods are stolen but not damaged no net loss to society occurs. Society has just as much wealth after the burglary as it did before. The wealth has simply been transferred from one individual to another. Both are members of the society, so there is no net loss. Others (e.g., Trumbell, 1990, and Cohen, 2000) exclude the private gains of criminals, and so view the theft as a loss. Likewise, Cohen would not count the suffering of people incarcerated in a cost-benefit analysis because they are criminals, while others would (Greenberg, 1990).

Even when people agree about the objectives, they may disagree about their priorities. In the case of drug treatment, one person may believe the social costs per gram of cocaine consumed, per serious crime, and per year of incarceration are $100, $10,000, and $25,000, respectively. Another might view drug use per se as less of a problem but believe that crime and incarceration carry hidden costs not reflected in budget-based estimates (e.g., fear of crime spurring middle class flight to the suburbs or the disenfranchisement of minority males by disproportionate rates of incarceration). Inasmuch as estimates of social costs reflect value statements, there is ample room for reasonable people to disagree about the relative costs of various outcomes and, hence, the relative desirability of various interventions.

A related problem stems from differences in opinion about the likelihood of different outcomes. Policy analyses of long-range social investments are fraught with uncertainties, many of which cannot be definitively characterized with objective, historical data. That is not a problem when there is a single decisionmaker. The methods allow and indeed even invite the inclusion of judgment in the form of "subjective probability assessments." But when many decisionmakers each have their own personal judgments about not only the likelihood of different outcomes but also the appropriate structuring of the problem, it is much harder for any single report or analysis to guide them collectively.

The result is that benefit-cost studies are sometimes performed from the perspective of a mythical "social planner," but they are read and judged by individuals with different agendas and different worldviews. A hypothetical early childhood intervention that is cost-justified by its effect on participants' crime rates a decade or more later when they are adults might not receive the support it "deserves" if the crime declines will bring rewards to the next generation of police commanders, rather than the current generation of social service agency heads, some of whom may not even think of crime prevention as the natural frame for evaluating the programs they sponsor.

Given these concerns, it is important to keep cost-benefit analysis, cost-savings analysis, and other forms of cost and outcome analysis in their place. They can provide valuable input to choosing among different programs, demonstrating a program's worth, improving programs, and replicating them. But they have their limitations. In any decision, some considerations can be resolved only through a decisionmaker's values and subjective judgment or through political interaction among stakeholders (Frank, 2000, Posner, 2000, and Richardson, 2000).

Chapter Three

ISSUES IN COST AND OUTCOME ANALYSIS OF EARLY CHILDHOOD INTERVENTION PROGRAMS

This chapter narrows the discussion of the methods described in the last chapter to the field of early childhood intervention programs. Following the framework of the scorecard presented in Chapter Two and in particular the row elements, we first outline some of the important issues related to measuring costs for these programs. Next, we describe the outcome domains that are relevant to early childhood programs, and how those outcomes are translated into program benefits (or costs avoided). The chapter closes with a discussion of specific issues in aggregation associated with comparing benefits and costs.

MEASURING COSTS OF EARLY CHILDHOOD INTERVENTION PROGRAMS

As discussed by Karoly et al. (1998), an extensive literature evaluates the impacts of early childhood intervention programs for participating children and their families. While evaluations of early childhood intervention programs have led to an established base of research focused on outcomes, there is less of a basis for assessing program costs. Information about program costs is often not reported in the evaluation literature and may not even be collected during the course of a demonstration project or larger-scale evaluation. As the discussion in Chapter Two conveys, however, cost information is an essential component of the types of cost and outcome analyses available to inform investment decisions in social service programs. Cost information may be used to evaluate the benefits versus costs to society of a given program or to guide innovation and improvement in program design.

Regardless of the goal of the cost and outcome analysis, the operational aspect of assembling the cost data is tantamount to filling in the cost element panel of the scorecard (Table 2.1). For each column in the scorecard, cost elements should be broken out by who bears various costs, when costs are incurred, and other aspects that would vary depending on the goal of the analysis. In the remainder of this section, we first discuss some of the general principles that guide cost measurement. We then discuss some of the details regarding the row elements and provide a hypothetical example of the cost elements in a scorecard.

Some General Principles of Cost Measurement

In measuring the costs of a program for the purposes of policy scorecard analysis, the goal is to enumerate the comprehensive set of resources forgone by all parties who might incur some loss as a result of the program. That is, the costs of a program are not entirely captured by the budget an agency uses to fund the program. Rather, a more comprehensive characterization of the costs of a program would capture the difference in resources required for a world *without* the program (the *baseline*) and the same world *with* the program, as discussed earlier in Chapter Two. This broader notion of costs allows for the fact that entities other than the agency—such as program participants and other members of society—might also incur some costs in a world with the program. It also recognizes that not all costs involve explicit expenses, but rather that some costs might take the form of in-kind resources devoted to the project, such as volunteer time or subsidized facilities.

This construct—program cost as the difference between total costs in a world with and without the program—highlights the importance of having a control (or comparison) group. Cost data from the control group serve as estimates of the costs of the world without the program, and data from the intervention group serve as estimates of the cost of the world with the program. If no cost data from a control group are comparable to the cost data from the intervention group, the estimates of the costs of the program will be fraught with considerably more uncertainty and error.

When collecting cost information for the control group, it is important that information be gathered for the same set of service providers as for the intervention group. This is essential to capture possible cost shifting (e.g., from one service provider or payor to another) or cost offsets (e.g., reduced use of services in one area as a result of increased services use in another). If cost information is more narrowly collected for the control group, it is possible to miss changes in the mix of services used or the total amount of services used as a result of the program (see, e.g., Foster and Bickman, 2000).

As discussed in Chapter Two, the objectives of the analysis dictate some of the particulars of cost estimation. For example, if the overarching goal is to compare the benefits and the costs of the program, then it is enough to estimate a single number or range of numbers (e.g., the cost is between $1.1 million and $1.3 million). However, suppose the objective is to estimate the cost of a similar program implemented somewhere else, or to use the cost estimates to guide a continuous quality improvement (CQI) effort. In these cases, it would be more useful to develop *cost estimating relations* (CERs), which estimate cost elements *as a function of the design of the program*. These relations generate various cost elements as a function of design variables, such as types of personnel who provide services, intensity of treatment, equipment and facilities required, and other potentially variable features of the program.

Types of Costs

There are various ways to categorize resources, but here we focus on some of the major categories that are likely to be particularly salient for early childhood programs. These categories help ensure comprehensive accounting of all resources that a program requires.

Cost analysts frequently categorize resources associated with program delivery into *personnel, equipment, facilities, and supplies/other*. Personnel includes all labor, e.g., social workers, nurses, secretaries, drivers, maintenance workers, and administrative personnel. Equipment includes durable items, such as office equipment (copiers, printers, computers, desks) and vehicles (automobiles, buses). Facilities includes land, office space, garage space, parking space, and maintenance sheds. Supplies/other includes consumable items, such as paper and ink for copiers and printers, gasoline for the vehicles, and coffee for the personnel. Utilities can be included in this category or broken out separately. In any particular study, if a category is too small (e.g., less than 5 percent of the total), the cost analyst may combine it with another. If a category is too large (e.g., more than 40 percent of the total), the cost analyst will split it into subcategories.

An important distinction in costs is between explicit expenses and in-kind resources. Obviously, costs that are billed need to be counted. It is also important to capture costs that accrue in the course of providing services but do not involve a monetary transfer. These likely will involve in-kind resources provided to the program from outside the agency, such as subsidized rent for facilities or meals provided by other government agencies.

Who pays for a resource is important. Cost analysts typically distinguish between *internal* and *external* costs. Internal costs fall on the agency that sponsors the program. External costs fall elsewhere. However, this distinction is often inadequate. Instead of distinguishing only between internal and external costs, one should distinguish costs (and benefits) by stakeholder. If there are a dozen stakeholders, there should be a dozen "who pays" categories.

For example, participants may bear certain costs to participate in the program. These would include the costs of transportation to appointments or lost wages from missed work. In the case of early childhood programs, it is especially relevant to consider costs borne not only by participating children, but also by their parents or caregivers, even when the latter group is not explicitly a focal point of the treatment program.

Another example of a cost that the agency providing services does not bear is the costs generated by referrals to other services. This is sometimes referred to as cost shifting and is important to capture in programs designed to increase use of other services. (Use of other services by providers outside the intervention may also decline.) For example, in the Elmira Prenatal/Early Infancy Project (PEIP)—a nurse home visiting intervention discussed more fully in Chapter Four (Olds et al., 1997)—part of the treatment provided by home visitors was to refer participants to other social services for which they might qualify. While greater use of these other social services did not impose a cost on the PEIP, it clearly raised costs for the other agencies that provided the additional services.

Collecting cost data for the control group (or baseline) for the same set of service providers as for the treatment group allows such cost shifting to be detected, though the analyst must give a priori thought to where cost shifting may occur and be sure to measure it. A comparison of cost data for the control group versus the treatment group will also reveal any cost offsets, whereby costs are reduced for services outside of the treatment program that are used by program participants.

Another way to categorize resources is to distinguish between consumable and nonconsumable items. A consumable item—such as paper or gasoline—is measured in units of quantity, such as reams or gallons. Nonconsumable items—such as facility space, durable goods, and personnel—are measured in units of quantity used per time unit—e.g., square-feet-months or person-years.

It is also frequently useful to distinguish between fixed and variable costs. Fixed costs are likely to be onetime costs, which often occur early in implementing a program. Examples of fixed costs are the costs of developing a curriculum or treatment protocol, and the costs of constructing facilities when they are not rented. The key feature of fixed costs is that they do not vary with the amount of time the program is in place. Variable costs are those that accrue in each time period the program operates, such as utility bills and payments to staff.

Cost analysts also distinguish between *investment* costs and *operating* costs. Investment costs are sometimes called *nonrecurring costs.* They are incurred to start a program or to increase its scale.[1] Often they pay for increases in nonconsumable resources such as vehicles or facilities. Operating costs are recurring costs; they must be paid each year to keep the program running. They are often assumed to be proportional to the inventories of nonconsumable resources on hand (e.g., salary plus benefits of an employee) or to the annual quantity of a resource consumed (the constants of proportionality are often called *cost factors*). Costs that have already occurred (or have already been contracted for) and cannot be recovered are *sunk* costs. These can correspond to resources that are on hand and cannot be sold, or in-kind contributions (e.g., volunteer labor, office space) which will only become available if the program is implemented. They should not influence one's decision whether to invest in the program, because they will be the same even if one does not invest.[2]

---

[1] In some studies it is useful to split nonrecurring costs into research and development (R&D) costs and investment costs.

Another noteworthy feature of costs is that they accrue over time and are likely to display variation over time. For instance, program costs might be high at the time of inception as the fixed costs of setting up facilities and training staff are born. Program costs might drop during a period when participants are screened or diagnosed, and then rise again during a "treatment" phase. It is useful to construct a variant of the scorecard whose rows are resource categories and whose columns are years. For consumable resources (those in the supplier/other category), each cell contains the amount of the resource consumed in that year of the program's operation. For nonconsumable resources, each cell contains the inventory of the resource on hand at the end of that year. This table is easier to construct than it may seem. Typically a program will start small and build capacity over time. So the analyst determines the resources needed by the mature program (say, in year five and beyond), and ramps up the resources over years one through four to achieve those levels.

In sum, cost accounting entails a number of categorizations. In filling in the cost elements of the scorecard, the analyst considers the various stakeholders, how costs accrue over time, whether the costs are in-kind or explicit, as well as a number of other considerations.

Capturing the Sources of Cost Variation

There might not be one simple answer to the question, "How much does this program cost?" The answer may depend on a number of factors, such as:

---

[2] There is an unfortunate temptation to let sunk costs affect one's decisions. "I'm going to hang onto that stock until it gets back to the price I paid for it." It is more profitable to base a decision to buy or hold or sell a stock on its future performance, not its past performance.

- From whose perspective are costs calculated—the participants, the government, society as a whole, etc.?
- In what geographic location or site is the program?
- What are specific design features of the program?

Capturing who pays for a particular resource enables one to calculate costs from different points of view. Social programs have many stakeholders, and a particular program may provide net benefits to some stakeholders while extracting net costs from others. It is not enough to calculate total net costs (and total net benefits) if the parties who pay the costs do not reap the benefits. A corollary to this is that costs may appear in more than one place in the cost model. That is, a cost element may be a cost to one party and a benefit to another, and hence will appear twice in the model (with opposite signs). When aggregated to society's perspective, these two would cancel out.

Even when program protocols are followed uniformly across locations, the program costs will likely vary by geographic location or site. This is because differences in costs result from such factors as the cost of rental space in a local area, whether a site is in an urban or rural area, and the relative wage rates of staff. For example, the transportation costs in rural areas, which often lack low-cost public transportation systems, might be considerably higher than those in urban areas. On the other hand, wages and rental prices are often lower in rural areas than in urban areas.

Another source of cost variation is the design features of the program. One's first inclination may be to take measures of the program's workload as the design variables. For an early childhood program, this might be the number of participants. However, additional design variables (e.g., capacity) are likely to be needed to portray costs fully. Often a program will be designed to have a given capacity, and some costs will be incurred whether or not the capacity is actually used. For instance, a group meeting requires a program staff member to be present no matter how well or poorly attended it may be. Omitting the capacity variables amounts to making an assumption about the utilization rate, which may have a strong influence on costs.

Another type of design feature that would generate variation in costs is the risk category of participants. Participants in different risk categories might utilize different types of services or require different lengths or intensities of treatment. Risk category may also yield differences in outcomes and hence benefits, and as a result the entire analysis could hinge on the distribution of participants across risk categories (Karoly et al., 1998).

Numerous additional design features might contribute to variability in program costs, such as whether the staff are medical doctors versus registered nurses or whether the participants are treated in a group or individual setting. While a complete list of design features would be too numerous to describe here, this discussion has suggested types of issues that need to be considered.

A Brief Hypothetical Example of Cost Elements and Data

To help fix the ideas we have discussed related to costs, we provide a brief example of some hypothetical cost elements and data in Table 3.1. Table 3.1 assumes a baseline and a new program, each of which provides some type of service to parents of young children. The cost elements and their values are completely fictional, but are a realistic representation of potential stakeholders, types of cost elements, and units of measure. In this table, we present hypothetical costs for the baseline program versus hypothetical costs for the new program. These are indicated by columns 2 and 3 in the table. The cost of the new program in this context—column 4—is the *difference* in costs between columns 3 and 2. That is, column 4 shows the incremental costs of the program over and above the baseline.

Cost elements are chosen to capture the resources employed by the program and who pays. We have chosen four hypothetical sets of stakeholders for whom there will be cost elements in this example: participants, the agency implementing the program, other agencies (that might provide services to which participants are referred), and the rest of society. We have indicated specific cost elements only for participants. These cost elements should represent all explicit, in-kind, and implicit resources the participants would incur when participating in the baseline program and the new program. These

Table 3.1

Hypothetical Examples of Cost Elements for Baseline Program and New Program

	Mean Hypothetical Costs		
Hypothetical Cost Elements (1)	Baseline Program (2)	New Program (3)	Difference Between (3) and (2)
Participants			
Number of visits per year	4	8	4
Time per visit	60 minutes	80 minutes	20 minutes
Wages per hour	$6.25	$6.25	0
Miles per trip to visit	15	15	0
Cost per mile	$0.20	$0.20	0
Copayments per visit	$5	$5	0
Paperwork time per year	35 minutes	50 minutes	15 minutes
Child care hours per year	0	8	8
Prescriptions filled per year	3	5	2
Agency Implementing Program			
Other Agencies			
Society as a Whole			

might include time expended and explicit cash outlays. Examples of cost elements in the table that might be assigned to participants are the time length of the visits, number of visits, visit copayments, paperwork time, number of prescriptions, and hours of child care (primarily when a parent is at additional appointments or meetings resulting from referrals).

It is important to measure cost elements in terms of resources, and only later price out the resources to obtain dollars. Costs should not be expressed directly in dollars, unless the resource inventories behind the dollars are unavailable. As was mentioned in Chapter Two, there are several reasons for this. These include variation in prices across locations, avoiding accounting assumptions regarding shared resources, and the need to substitute for resources not available in another site.

In Table 3.1, this is demonstrated in several cost elements. Transportation, for instance, is expressed in the number of trips rather than dollars, and the miles per trip and cost per mile are indicated in separate cost elements. This way, sites where participants use different modes of transportation, such as subway, bus, cab, or the participant's car, can account for the differences in time costs and cash outlays inherent in those modes of transportation. Similarly, the time per visit is expressed in minutes. Sites that serve participants who work would value this time differently from sites whose participants are largely out of the labor force.

As discussed in Chapter Two, scorecard entries will ideally include information that characterized their statistical uncertainty. While not shown explicitly in this hypothetical table, cost ranges or confidence intervals could be included as well as expected values. The degree of statistical uncertainty surrounding cost estimates would be important to consider when comparing costs across programs.

As a final point in discussing costs, note that gathering data for the analysis described here is itself a cost. In addition to the resources consumed by the analysis team, note that the analysis is likely to impose costs on the program itself. The program staff likely will be required to provide or collect data, which will require additional time, training, and perhaps even computers or other equipment.

OUTCOME DOMAINS AND MEASURING BENEFITS

As noted above, there is a long history and well-developed methodology for measuring the impacts of early childhood intervention programs on participating children and their families. Since the 1960s, a wide array of smaller- and larger-scale early intervention programs have been implemented and formally evaluated, often with experimental designs to allow comparison of outcomes for program participants versus a randomly assigned control group.[3] These evaluations generate specific measures of program impacts for particular individuals at given points in time, either when program services are delivered or after the intervention program has ended.

---

[3] While the randomized control trial remains the gold standard for evaluating social service delivery programs, some evaluations adopt quasiexperimental designs using matched comparison groups as controls. The experimental and quasiexperimental early intervention evaluation literature has been synthesized in a number of comprehensive reviews. For recent examples, see Barnett (1995), Yoshikawa (1995), Guralnick (1997), Reynolds et al. (1997), Karoly et al. (1998), and Currie (forthcoming). Regression analysis and related methods are another set of tools that can provide insights into service delivery questions, such as whether a program improves outcomes of participants or whether different populations realize different outcomes from a program (see, for example, Currie and Thomas, 1995, or NICHD Early Child Care Research Network, 1997). Also, see Hargreaves et al. (1998), Chapter 9, for a discussion of the use of regression techniques in cost and outcome methods.

Just as program costs are measured net of a baseline without the program being evaluated, the outcomes of early intervention programs are net impacts (i.e., the same outcomes are measured for both the treatment and control/comparison groups for the same period of follow-up, and the program effects measure the difference between the two groups). Typically, these program impacts are measured in quantities other than those denominated in dollars. The challenge for cost and benefit analysis is to translate the beneficial effects of early intervention programs measured in such units as IQ points, years of special education, months employed, or counts of juvenile crimes into dollar values that can then be compared with program costs. The remainder of this section considers the types of program impacts typically included in evaluations of early intervention programs and the approaches available for translating these outcomes into dollar benefits.

Measuring the Impact of Early Childhood Intervention Programs

Targeted early intervention programs can be viewed as sharing a common aim: to improve child health and development by providing socioeconomically disadvantaged children and their families with various services and social supports during part or all of the period of early childhood (Karoly et al., 1998). Despite this common aim, considerable variation occurs in early intervention program objectives and designs and in the associated services and supports provided to meet the program goals. Likewise, program evaluations are not uniform in the outcome measures collected. Instead, resource constraints for data collection and other factors limit most evaluations to capturing only a subset of measures that reflect the domains where the program is expected to have an impact, whether for the focal child or for their parents and other caregivers.[4]

---

[4] While program evaluations do not always collect the same set of outcome measures, for those measures conceptually similar to those collected in other evaluations, it is often desirable to use common measures so that comparisons can be made. For

Table 3.2 illustrates the range of measures of early intervention program impacts in four broad domains: emotional and cognitive development, education, economic well-being (e.g., public assistance receipt, income, crime), and health. Within each domain, we list some of the most frequently used measures in early intervention studies, either for participating children or for their parents and other caregivers.[5] (In support of the discussion in the next subsection, italics are used in the table to indicate which of these outcomes are most readily translated into dollar values.) The specific measures in Table 3.2 are intended to illustrate the types of outcomes measured in each domain, rather than reflecting the full range of measures used in the evaluation literature. We discuss each of these domains in turn, as well as some more general measurement issues common across domains.

Emotional and Cognitive Development. Given the goal of early intervention to enhance child development, most early intervention evaluations include measures in this domain, either for participating children or their parents and other caregivers. For children, the measures include scores on batteries that measure socioemotional development or behavioral problems, as well as cognitive development—typically IQ scores. For parents, scales are used to measure aspects of the parent's role in the child's development, such as the nature of the parent-child relationship and the quality of the home environment. The specific scales and tests used are selected to be age-appropriate (whether administered to a parent or child) and to reflect the specific objectives of the program being studied. To select measures that are reliable and valid, with well-known psychometric properties, many interventions often use the same specific scales or test batteries, such as the Stanford-Binet or Wechsler intelligence tests to measure IQ, or the HOME Inventory to assess parental caregiving and the home environment.

---

example, in the early childhood intervention literature, certain test batteries or scales are often used to measure cognitive or behavioral development, and information on labor market outcomes or income can be collected in a uniform way. A review of previous evaluations can aid in the design of data collection protocols so that the outcomes from the program under consideration can be compared with similar programs that have also been evaluated.

[5] For additional detail, with examples of outcomes measured for specific studies, see Karoly et al. (1998) and the other literature reviews referenced above.

Table 3.2

Early Childhood Intervention Program Outcome Domains and Illustrative Measures

	Illustrative Measures for:	
Outcome Domain	Child	Parent/Caregiver
Emotional and cognitive development	Socioemotional and behavioral scores IQ test scores Teacher's ratings	Quality of parent-child relationship Quality of home environment
Education	Achievement test scores Grades *Grade progression (repetition)* *Participation in special education* *Educational attainment*	*Educational attainment*
Public assistance receipt, income, crime	*Receipt of public assistance* *Employment* *Earnings/income* *Criminal activity* *Contact with criminal justice system*	*Receipt of public assistance* *Employment* *Earnings/income* *Criminal activity* *Contact with criminal justice system*
Health	Physical and mental health status *Child abuse and neglect* Substance abuse Fertility control *Emergency room visits* *Other health care use*	Physical and mental health status Family violence Substance abuse Fertility control

NOTE: *Italics* indicate measures more easily expressed in dollar terms.

Education. Another common aim of early intervention programs is to improve school readiness and subsequent school performance. Consequently, a great deal of interest has arisen in tracking educational outcomes for program participants versus those in the control or comparison group. Prior to school entry, few direct measures of school readiness exist although researchers often consider measures of cognitive development and socioemotional regulation and control as relevant indicators. For school-age children, evaluations typically measure scores on achievement tests in reading, math, or other subjects. Achievement test scores at older ages are included in longer-term evaluations as well. Longer-term follow-up also allows measurement of educational outcomes relevant at older ages, such as grade progression (or alternatively grade repetition), use of special education, high school completion, and eventual educational attainment. In some cases, early intervention programs are designed to improve educational outcomes for parents as well, so educational attainment is also measured for them.

Economic Well-Being. Early intervention programs may also affect other areas of functioning during adolescence and adulthood. If early intervention programs improve socioemotional development and educational performance, those gains may translate into improved economic well-being. With longer-term follow-up, for example, some programs have been evaluated in terms of their impact on economic outcomes such as dependence on social welfare programs (e.g., use of public assistance or "welfare," Food Stamps, or Medicaid) and labor market performance or economic success (e.g., employment rates, occupational status, earnings, income, poverty status). Another area of assessment is involvement in criminal activity, either by directly measuring specific crimes committed or by quantifying contact with the criminal justice system (e.g., arrests, convictions). While program evaluations typically consider these outcomes for participating children as they make the transition to adulthood, some programs have assessed parents and other caregivers in this domain using similar measures.

Health. This final category captures the expectation that early intervention programs may affect health outcomes, broadly defined to include aspects of health status and health care use. In addition to evaluating the impact of early intervention on general physical or mental health status, some programs consider more specific areas of health, such as the incidence of child abuse and neglect, perceived quality of life, family violence and substance abuse, impairment, and fertility control (e.g., the timing and spacing of births). Health care use may also be affected by early intervention programs, with some programs focusing on costly emergency room visits, as well as other forms of health care use (e.g., hospitalizations or use of specific health care services). While many program evaluations focus on these measures for participating children, either at younger or older ages, these measures may also be assessed for parents and other caregivers when early intervention services are designed to affect their health status or health care use.

Across these four domains, a number of common measurement issues arise. A first concern is whether to focus outcome measurement on the participating child or the parent and other caregivers, and for each of these potential beneficiaries of program services, whether the intervention impacts can be captured in the short run or the long run. As indicated in Table 3.2, early intervention programs may benefit not only the children they serve but also their parents or caregivers. The first generation of early intervention programs and their associated evaluations focused on child outcomes (see, for example, the studies cited in Karoly et al., 1998). With a growing recognition of the importance of the family and home environment in the early years of life and of the potential for programs to impact parental outcomes, program services and evaluations have incorporated the parental side of the equation as well. If a program can be expected to affect parental outcomes, many of those outcomes listed in Table 3.2 can be captured in a short-term evaluation, as they may be measured during the period of service delivery or soon after the program ends. In contrast, many of the outcomes in Table 3.2 listed for children cannot be directly assessed without follow-up that extends many years, if not multiple decades, beyond the period of program delivery. Such long-term follow-up requires a significant commitment of resources to execute as well as to minimize the biases associated with attrition in longitudinal studies.[6]

Another methodological concern is whether measures should capture contemporaneous outcomes or a longer history of a given outcome. This is particularly relevant for evaluations that include long-term follow-up. Consider the case of public assistance utilization, either by the participating child's family during childhood or by the child when the child reaches adulthood and forms a household of his or her own. During any given assessment, either during the intervention or in a subsequent follow-up, it is possible to collect information on current program utilization (i.e., whether the individual is currently receiving support). An alternative is to collect data on a partial or complete history of program participation during the interval between the present and the last point of data collection (or even prior to the baseline). While the intervening history clearly requires more effort to collect, it provides the information necessary to value a continuous sequence of potential benefits. A reduction in public assistance utilization in each year of a 10-year horizon will clearly translate into greater savings to government than will a reduction in utilization for the final year of that horizon (e.g., the tenth year). This consideration is relevant for many of the outcomes listed in Table 3.2, including measures of educational outcomes (e.g., grade repetition, special education use), economic outcomes in addition to use of social welfare programs (e.g., employment, earnings, income, criminal activity, and criminal justice system contact), and health outcomes (e.g., health care utilization).

---

[6] It is important to standardize the period of follow-up or future projection if programs are to be strictly compared in terms of costs and benefits. Otherwise, the program with a longer follow-up or with projections further into the future will likely be favored on cost-benefit terms.

A final measurement issue is the method of collecting the specific outcome indicators. During the period of program intervention, the measures listed in Table 3.2 (and others in the four domains not listed) are typically collected through some form of interaction with the study participants (those receiving the treatment as well as controls). Survey questionnaires, test batteries, direct observations, and program administrative records may be appropriate depending on the specific outcome of interest. Once the intervention has ended, continued assessment may require continued personal interaction with study participants or possibly the reliance on external sources of information, such as administrative records. For example, with the proper human subjects consent procedures, information on criminal activity (e.g., arrests, incarcerations) may be collected by interviewing participants during a follow-up or by tracking activity recorded by the criminal justice system. Administrative data can be useful for collecting information on other outcomes, such as school performance, participation in social welfare programs, and employment outcomes.

Administrative data have several advantages. They may be free of various reporting biases and may result in lower rates of missing data (or cases lost due to nonresponse). This is especially true for longer-term follow-up when respondents may have difficulty with long-term recall of specific events (e.g., a monthly employment history) or may not be even available for an interview because they cannot be located. However, tracking outcomes through administrative sources requires advanced planning to secure the necessary permissions from study participants. Administrative data are often not released due to concerns about protecting individual privacy, and individuals may still be lost to follow-up when they cannot be tracked across administrative boundaries (e.g., state borders).

Translating Program Impacts into Dollar Benefits

Once a formal program evaluation has measured the impact of an early intervention program using one or more of the measures listed in Table 3.2, many of the analysis methods reviewed in Chapter Two require that the analyst convert that outcome to a monetary value. The process of expressing the benefits in dollar terms, or "monetizing" the program impacts, is easier for some of the outcomes listed in Table 3.2 than for others. This reality is illustrated in Table 3.2 by denoting those outcomes that are most readily monetized in italics. Those outcomes not in italics may still be expressed in dollar terms but only with less reliable benefit-cost estimates or by virtue of more heroic assumptions.

The economic outcomes listed in Table 3.2 are among those that are the easiest to monetize, whether the program impact is lower public assistance utilization and the benefit is reduced outlays by local, state, or federal governments or the program impact is more months spent employed and the benefit is higher taxes paid.[7] To illustrate, consider an early intervention program evaluation, shich shows that at the age 15 follow-up, the families of children who participated in the program used 20 fewer months of public assistance benefits over the past 10 years than did families in the control group. If each month of benefits costs taxpayers $500 (including both cash payments and administrative costs), this early intervention program would lead to dollar savings to government of $10,000. After spreading those savings over each relevant age (from six to 15) and discounting to a specific time period (e.g., birth) using a specific discount rate (e.g., 4 percent), the NPV of the savings to government could be calculated ($6,410 in this case). A similar process can be followed for each of the italicized outcomes shown in Table 3.2.[8]

---

[7] These are benefits from the point of view of the government, and we adopt this point of view for illustration only. The analyst should be prepared to present costs and benefits from the point of view of any stakeholder. For example, from the point of view of society as a whole, taxes are not a benefit but a transfer payment, and one should use incremental income instead.

Other outcomes listed in Table 3.2 cannot be translated into dollar values with the same ease. For example, many early intervention programs demonstrate short-term and long-term gains in cognitive measures, such as IQ or achievement test scores. This impact is difficult to translate into a dollar value. However, if these cognitive benefits lead to improved educational and economic outcomes, then valuation of outcomes in these collateral domains may capture, at least in part, some of the benefits of better cognitive outcomes.

The process of assigning an economic value to a given program impact is not always uncontroversial. One outcome that invites differences of opinion is the value to society of the reduction in criminal activity stemming from early intervention. As discussed in Chapter Two, while most experts agree on the value of the tangible costs associated with criminal acts based on empirical evidence (e.g., costs for property loss, medical expenses, lost income due to injury), there is less agreement over the value to assign the intangible costs (e.g., pain and suffering of crime victims). Different methods of valuing pain and suffering can lead to widely different estimates of the intangible costs of crime. For instance, Klaus (1994) estimates the cost of a rape to be $234, whereas Miller et al. (1996) use a figure of $5,100. Based on personal experience, some audiences believe a figure in the $5,000 range is much too low. This type of controversy may affect other areas of program impacts, particularly when empirical evidence regarding economic values is weak or nonexistent.

In some cases, it is possible to assign benefits beyond the period of direct observation. For example, improvements in educational attainment can be associated with an entire earnings profile from young adulthood to age 65 based on other studies of earnings trajectories in the literature (for an example, see Barnett, 1993). On the basis of criminal activity through adolescence or early adulthood, the individual's future criminal "career" in adulthood can be forecast

---

[8] For more detail on these types of calculations, see the cost-benefit studies cited in Chapter Four.

(see Karoly et al., 1998, for an example). Although such forecasts introduce additional uncertainty into the benefit calculations, they do help overcome the limits of follow-up periods that end in early adulthood when the economic benefits for participants in early intervention programs may just be beginning to be realized.

When intermediate impacts (e.g., educational attainment) are used to value longer-term impacts (e.g., earnings), it is important to avoid double counting program benefits. In some cases, the intermediate outcome may generate benefits in and of itself, in addition to providing information to project benefits for a longer-term but unobserved outcome. In the case of educational attainment, if an early childhood program increases years of schooling for the treatment group compared with the control group, educational costs actually *increase* because of the additional time spent in school. At the same time, the higher educational attainment can be used to project earnings gains throughout adulthood compared with the trajectory that would be experienced with a lower level of attainment. However, if actual earnings are observed for any period beyond the intervention, the projected earnings should not be counted for the same age span.

COMPARING COSTS AND BENEFITS

In comparing the costs and benefits of an early childhood intervention program, two critical issues are the following:

- Who pays the costs versus who realizes the benefits?
- What is the decision rule for selecting the best alternative?

These two issues are related. We discussed the first issue earlier in pointing out that various costs are borne by different stakeholders, such as children, parents, government, and society as a whole. The benefits could be realized by one party—such as the children—while the costs are paid by another party—say, the government.

The second issue helps resolve this quandary. As discussed in Chapter Two, various decision rules might be specified that would yield different answers to the question of whether the benefits outweigh the costs and for whom. In *cost-savings analysis*, the costs of a program to government are compared to the savings of a program to government. If the latter outweigh the former, then the program pays for itself from the perspective of government. In *cost-benefit analysis*, the costs of the program borne by all of society—including participants, government, and others—are compared to the total benefits accrued to any of the parties. This calculus is indifferent to who pays and who benefits.

Chapter Two noted some specific methodological issues associated with these various approaches, including choosing a discount rate, accounting for scenario uncertainty, and capturing statistical uncertainty. An additional challenge in the comparison of costs and benefits likely to be particularly relevant for early childhood intervention programs is the fact that they may accumulate at different rates. These programs typically intervene briefly in the early years of a child's life. In contrast, the benefits may take years to accumulate, as the child's outcomes in such areas as high school graduation, adult employment, and public assistance participation become apparent. This creates a potential temporal mismatch between the payment of costs and the realization of benefits, even if the measure of merit considers only costs and benefits to government. This is because the government—i.e., the taxpayers—that pays for the program might not be the same government (taxpayers) that reaps the benefits two decades later when the treated children enter adulthood.

As discussed further in the next chapter, Karoly et al. (1998) demonstrate that the costs of the Perry Preschool program take two years to accumulate compared to the benefits, which accumulate to the level of costs after nearly two decades. In contrast, another program reviewed in the next chapter, the nurse-home visiting model known as the Elmira Prenatal/Early Infancy Project, generates benefits earlier in the life course because of changes in the parents' behavior (specifically, the mother). In that case, the benefits accumulate more rapidly and are realized at a level that exceeds program costs shortly after the two-year intervention ends.

Another challenge for these tools is the conservative nature of most estimates of program benefits. Due to the limitations of placing an economic value on the benefits of early intervention, most cost-benefit studies of these programs are likely to understate the benefit side of the equation for two reasons.[9] First, many of the benefits of early intervention programs may not even be measured as part of the evaluation. This may stem from resource constraints that limit the number of measures collected or because some measures are more difficult to collect. For example, some early intervention programs may produce spillover benefits for other siblings (e.g., as a result of improved parenting or better economic situation of the family), or may lead to spillover benefits for other children in the child's community (e.g., at the same school or in the same neighborhood). Measuring these types of potential spillover benefits is more costly. If these outcomes are not included in the evaluation, it is even more difficult to incorporate them into a cost-benefit calculation.

Second, many of the benefits captured in an evaluation cannot be expressed in monetary terms, either as benefits to program participants or to the rest of society. As illustrated in Table 3.2 (and in the specific studies reviewed in Chapter Four), only a subset of the outcomes that may be affected by an early intervention program can be readily expressed in monetary terms. In other cases, the assumptions needed to assign a monetary value to a given outcome are so heroic that it is preferable to err on the side of undervaluing a program's benefits.[10] To the extent that cost data are easier to collect and less subject to under- or overestimation, cost-benefit calculations for early intervention programs will likely err on the side of being conservative.[11]

---

[9] This assumes that the dollar values assigned to those program benefits that can be monetized are not biased upward or downward.

[10] The use of the scorecard still allows the decisionmaker to account for benefits that are not monetized and to use his or her own subjective weights in valuing those outcomes. See Sen (2000) for further discussion of this issue.

[11] This conclusion rests on the assumption that cost data are less likely to be underestimated or overestimated. This may be reasonable for those costs directly associated with service provision. However, indirect costs may be equally hard to measure or estimate as some of the benefits listed in Table 3.2. Data collection constraints may also result in underestimation of program costs if not all areas of cost are measured (e.g., cost shifting). However, given that costs are typically incurred during a fixed interval of program provision, while benefits may accumulate indefinitely into the future, the inability to capture the (discounted) monetary value of long-run benefits in certain domains is likely to outweigh the short-term costs that are underestimated. Because this conclusion is by no means universal, whether net benefits are under- or overestimated needs to be assessed on a case-by-case basis.

Chapter Four

BENEFIT-COST FINDINGS FOR EARLY CHILDHOOD INTERVENTION PROGRAMS

In the past several decades, a number of early childhood intervention programs have been rigorously evaluated to assess their impact on participating children and their families (see, e.g., the studies reviewed in Barnett, 1995; Currie, forthcoming; Karoly et al., 1998; Lazar and Darlington, 1982; Reynolds et al., 1997; White, 1985). While this literature is extensive and provides strong evidence that early intervention programs can produce significant short-run and long-run benefits for participants, only a handful of programs have been subject to a formal cost-benefit analysis.

To illustrate the cost and outcome methods discussed in Chapters Two and Three, in this chapter, we review the findings from three early intervention programs that have been evaluated in terms of program costs and benefits.[1] In each case, we provide a brief summary of the early intervention program and evaluation findings, as well as the results for the cost-benefit analysis.[2] A final section compares the cost-benefit findings across the programs and the implications for cost-benefit analysis of early childhood intervention programs more generally.

---

[1] Cost-benefit analyses are expected to be available soon for other programs in addition to those we review in this chapter. For example, a cost-benefit analysis is under way for the Carolina Abecedarian program based on follow-up information through age 21 for the participants in this center-based early intervention program (Campbell and Ramey, 1994). Also, Currie (forthcoming) provides a "back-of-the-envelope" cost-benefit calculation for the Head Start program based on both short-term and long-term benefits generated by the program. These calculations suggest that even considering only a subset of the short- and medium-term benefits, Head Start already pays back much of the program costs. With modest-size long-term benefits, the full benefits of Head Start would likely more than pay back the program costs although more in-depth benefit and cost analysis is required to confirm this rough calculation.

THE PERRY PRESCHOOL PROGRAM

The High/Scope Perry Preschool program is perhaps the best-known center-based early intervention program, in part because of the long-running experimental assessment that has demonstrated the program's effectiveness (Schweinhart et al., 1993). This small-scale, model program served 58 African American children between 1962 and 1967 in Ypsilanti, Michigan, beginning at age three for two years of program services or age four for one year. Another 65 children were in the randomly assigned control group. Children were selected from among low socioeconomic status (SES) families where the child scored less than 85 on a standard IQ test.

Those in the Perry Preschool program attended 2.5-hour center-based classes and 90-minute teacher home visits between October and May of each year. The program is known for the high quality of the teaching staff and the low pupil-teacher ratio, as well as the richness of the curriculum. Both the participants and the control group have been followed through age 27.

Program Benefits

Table 4.1 summarizes the impact of the Perry Preschool program in four key domains: emotional and cognitive development, education, economic well-being, and health. In this case, all measured outcomes focus on the children in the treatment group compared with the control group. As with other early intervention studies of the era, the first outcomes measured were changes in IQ. At the end of the program intervention, children in the preschool program had IQ scores that exceeded the control group by 12 points. The positive IQ

---

[2] The next three sections of this chapter draw heavily on Karoly et al. (1998) and Karoly (forthcoming).

Table 4.1

Measured Outcomes and Results for Selected Early Intervention Programs

Program and Subject	Cognitive/ Emotional Development	Education	Economic Well-Being	Health
High/ Scope Perry Pre-school Project: Child	IQ at age 5 (SB): E>C E=94.9, C=83.5 IQ at age 7 (SB): E>C E=91.7, C=87.1 IQ at age 8 (SB): E=C E=88.1, C=86.9 IQ at age 14 (W): E=C E=81.0, C=80.7	Achievement tests at age 9: E>C Achievement at age 14: E>C High school GPA at age 19: E>C Time in special ed. through age 19 (% of years): E<C E=16%, C=28% Years in educable mentally impaired programs through age 27: E<C E=1.1, C=2.8 Years retained in grade through age 27: E=C E=0.5, C=0.7 High school graduation rate by age 27: E>C E=66%, C=45% Postsecondary ed. credits by age 27: E=C E=33%, C=28%	Ever arrested by age 27: E<C E=57%, C=69% Lifetime arrests through age 27: E<C E=2.3, C=4.6 Employment rate at age 19: E>C E=50%, C=32% Employment rate at age 27: E=C E=71%, C=59% Monthly earnings at age 27 (1993 $): E>C E=$1,219, C=$766 Received public assistance at age 27: E<C E=15%, C=32% Received public assistance in last 10 years at age 27: E<C E=59%, C=80%	Teen pregnancies per 100 females through age 19: E=C E=68, C=117

Table 4.1—continued

Program and Subject	Cognitive/ Emotional Development	Education	Economic Well-Being	Health
Elmira PEIP[a]: Child	IQ at age 3 (SB): E=C E=103.6, C=102.0 IQ at age 4 (SB): E=C E=111.5, C=108.9	NM	Arrests through age 15: E<C (HR) E=0.24, C=0.53 Convictions through age 15: E=C E=0.13, C=0.18	ER visits through ages 25–50 mos.: E<C E=1.0, C=1.5 Hospital days through ages 25–50 mos.: E>C E=0.54, C=0.30
Elmira PEIP[a]: Mother	HOME Inventory at 46 mos.: E=C Reports of child abuse and neglect through age 15: E<C E=0.29, C=0.54	Years of ed. at age 4: E=C E=11.4, C=11.1	Mos. employed through age 15: E=C E=96.4, C=89.7 Mos. on AFDC through age 15: E<C (HR) E=60.4, C=90.3 Mos. on Food Stamps through age 15: E<C (HR) E=46.7, C=83.5 Mos. on Medicaid through age 15: E=C E=61.8, C=70.0 Arrests through age 15: E<C (HR) E=0.18, C=0.58 Convictions through age 15: E<C (HR) E=0.06, C=0.28 Jail days through age 15: E<C (HR) E=0.04, C=1.11	Subsequent pregnancies through age 15: E<C (HR) E=1.5, C=2.2 Subsequent births through age 15: E<C (HR) E=1.1, C=1.6 Mos. between first and second birth through age 15: E>C (HR) E=64.8, C=37.3 Substance use impairments through age 15: E<C (HR) E=0.41, C=0.73

Table 4.1—continued

Program and Subject	Cognitive/ Emotional Development	Education	Economic Well-Being	Health
Chicago CPC[b]: Child	Teacher ratings of school adjustment at age 9: E=C	Achievement tests at age 9: E>C Achievement tests at age 14: E>C Special ed. by age 9: E=C E=8%, C=9% Special ed. (yrs.) through age 14: E<C E=0.6, C=0.9 Special ed. (yrs.) through age 18: E<C E=0.7, C=1.5 Grade retention by age 9: E<C E=19%, C=26% Grade retention by age 14: E<C E=25%, C=37% HS graduation rate by age 20: E>C E=49%, C=39%	Delinquency rate through ages 13–14: E<C Delinquency rate through ages 15–16: E=C Juvenile court petitions through age 17: E<C E=16%, C=26%	NM
Chicago CPC[b]: Parent	Parental involvement in school at age 9: E>C	NM	NM	NM

SOURCE: Karoly et al. (1998), Tables 2.2 and 2.3 with updated information for the Chicago CPC program based on Reynolds et al. (2000).

NOTE: Statistically insignificant results are designated by E = C; results significant at the 0.05level or better are designated by E>C or E<C. NM = not measured; HS = high school; E = experimental, C = control. For IQ tests: SB = Stanford Binet; W = Wechsler Intelligence Scale for Children. The HOME Inventory assesses aspects of parental care giving and characteristics of the physical home environment.

[a] Results are for full sample unless otherwise indicated. "Higher-risk" (HR) only is defined as single mothers with low SES for child results through age 15.

[b] Results for preschool treatment group versus no preschool comparison group.

effect for program participants began to decline after school entry, disappearing by second grade (age eight) (Schweinhart and Weikart, 1980).

These early positive IQ effects were followed by improved academic achievement even after differences in IQ between the groups ceased to be statistically significant. For instance, achievement test scores for program participants remained significantly higher than the control group through age 14. Preschool participants had better grades and were more likely to graduate from high school; at age 28, there were no differences in postsecondary education participation, however (Schweinhart et al., 1993). The differences in rates of special education and grade repetition by age 27 were in the expected direction and statistically significant for the former measure.

At the last follow-up at age 27, other lasting differences were evident as well in employment, welfare, and crime outcomes (Schweinhart et al., 1993; Barnett, 1993). For instance, by age 27, program participants had significantly lower rates of current and past welfare utilization (i.e., AFDC, Food Stamps, and so on). Lifetime criminal activity—both incidence and severity—was also significantly lower. Employment rates and earnings for program participants were higher, although the employment rate difference was not statistically significant. Health effects, in contrast, were not as strong. The difference in the teen pregnancy rate by age 19 was large in absolute terms (68 per 100 females for the treatment group versus 117 per 100 females for the controls) but only marginally significant given the small sample size ($p = .08$).[3] Other behaviors include a statistically significant higher rate of marriage by age 27 among women participants in the preschool program (Schweinhart et al., 1993).

Cost-Benefit Analysis

Along with the extensive evaluation of the outcomes of the Perry Preschool program, a number of cost-benefit analyses have been conducted. Those based on the data through the age 27 follow-up include Barnett's (1993) original analysis and a reanalysis by Karoly et al. (1998). Barnett's estimates, consistent with early cost-benefit assessments of the program, indicate that benefits to society exceed program costs by a factor of more than seven to one. The largest component of benefits is from reductions in crime, a large fraction of which is the estimated reductions in the intangible losses to victims of crime over the lifetime of the program participants. Other large savings components include taxes recovered over participants' lifetimes due to higher earnings, and reduced K through 12 education costs.

---

[3] The birth rates are calculated based on the total number of pregnancies and live births per woman in either treatment or control groups by age 19. The 24 women in the control group had a total of 28 births in contrast to 17 births for the 25 women in the treatment group.

Karoly et al. (1998) use the results from Barnett's (1993) analysis but adjust his figures to 1996 dollars and rediscount benefits and costs to the birth of the focal child using a 4 percent real discount rate to be consistent with the method adopted for the cost-benefit analysis of the Elmira PEIP reported below. Like Barnett's (1993) approach, Karoly et al. (1998) express savings to government in monetary terms from the following outcomes observed for participating children compared with the controls:

- Reduced use of special education and fewer years of grade retention (net of increased education costs due to greater educational achievement) through age 27.
- Increased taxes from higher employment projected through age 65 based on employment and earnings data at age 27.
- Less time spent on welfare projected through age 65 based on welfare utilization observed through age 27.
- Reduced criminal justice system costs projected for their lifetime based on outcomes observed through age 27.

(These benefits are among those cited in Table 4.1 for the Perry Preschool program but do not include benefits in domains that are harder to express in monetary terms, such as higher IQ or achievement test scores.)

In addition to the savings to government, the cost-benefit analysis by Karoly et al. (1998) quantified benefits to the rest of society in two areas: the increase in net income for program participants stemming from higher work effort and earnings (net of reductions in welfare payments) and the reductions in the tangible costs associated with criminal activity (i.e., property loss, medical expenses, and income lost while injured). Barnett's (1993) analysis differed from this approach in that both tangible and intangible crime benefits were incorporated into the analysis, where the latter includes the value of reductions in pain and suffering associated with the reduced criminal activity.

The present discounted value of government savings and benefits to the rest of society can be compared with program costs. Barnett (1993) reports that the Perry Preschool program cost $12,356 in 1992 on average per child.[4] After inflating the costs to 1996 dollars to account for inflation and after discounting to birth using a 4 percent real discount rate, Karoly et al. (1993) report that the program costs $12,148 on average per participating child.

Table 4.2 summarizes the results of the cost-benefit analysis reported by Karoly et al. (1998), showing program costs, then the component elements of savings to government, and finally the components of savings for the rest of society. The net benefits are shown in the final row of the table. As shown, all benefits accrue from changes in the child's behavior. (This contrasts with the results for the PEIP, where some benefits are due to improvements in the mother's outcomes.)

As illustrated in Table 4.2, the Perry Preschool program produces savings to government more than twice the program costs ($25,437 versus $12,148), and a similar ratio results for the monetary benefits to the rest of society.[5] Consequently, the total benefits (savings to government plus benefits to the rest of society) are estimated to exceed program costs by a factor of four to one, with net benefits of $37,824 per child served.[6] The largest component of benefits mea

---

[4] This is a weighted average that accounts for the fact that about 20 percent of participants attended only one year of the two-year program (Barnett, 1993).

[5] To account for statistical uncertainty, Karoly et al. (1998) also calculate a confidence interval for the estimate of government savings and show that, while the error bands are large, the likely range of net savings to government is still positive.

[6] Barnett (1993) estimates a ratio of total benefits to costs of seven to one stemming from the valuation of certain kinds of intangible benefits to the rest of society from reductions in criminal activity (e.g., reduced pain and suffering experienced by crime victims).

Table 4.2

Costs and Benefits: Perry Preschool Program

	Dollars per Child		
	Due to Mother	Due to Child	Total
Program cost			12,148
Savings to government			25,437
Reduction in education services	*	6,365	
Reduction in health services	*	*	
Taxes from increased employment	*	6,566	
Reduction in welfare cost	*	2,310	
Reduction in criminal justice cost	*	10,195	
Additional monetary benefits			24,535
Increase in participant income net of welfare loss	*	13,846	
Reduction in tangible losses to crime victims	*	10,690	
Total benefits			49,972
Net benefits			37,824

SOURCE: Karoly et al. (1998), Tables 3.6 and 3.7.

NOTE: * = not measured. All amounts are in 1996 dollars and are the NPV of amounts over time where future values are discounted to the birth of the participating child, using a 4 percent annual real discount rate.

sured (about 40 percent) is the savings to government and benefits to the rest of society from the reduction in criminal activity for Perry Preschool program participants. Another significant component is the increased net income for participants in the program, although this component would not be immediately available to the government to pay for the program (unless these gains are taxed away). Savings to government from lower educational expenses and increased taxes each account for about 13 percent of the benefits generated.

THE ELMIRA PRENATAL/EARLY INFANCY PROJECT (PEIP)

The home visiting model is the second major paradigm in the early intervention literature and the Elmira Prenatal/Early Infancy Project (PEIP) is among the best-known in this class, in part again, because of the long-term experimental evaluation of the program (Olds et al., 1997).[7] The PEIP provided nurse home visits to a mostly white sample of first-time mothers in Elmira, New York, between 1978 and 1980. The program targeted higher-risk women (e.g., pregnant teenagers, low SES, single-parent households) although the program was open to all first-time mothers who asked to participate.

Through a series of prenatal visits, the trained nurse home visitors worked with the mother to improve her pregnancy outcome. After the child's birth, the nurse worked with the mother to improve her parenting skills and increase her economic self-sufficiency by linking her with various social services. The visits continued until the child was age two. On average, the nurses completed nine visits during pregnancy and 23 visits from birth to age two. Participants in the Elmira randomized control trial (300 total in the treatment and control groups) have been followed through age 15, with a focus on outcomes both for the mother and the focal child.[8] For purposes of analyzing the long-term follow-up results of the Elmira PEIP, Olds et al. (1997) report results for the full experimental group, as well as a higher-risk subsample. This latter group consists of women who, at the time of enrollment in the study, were unmarried and had low SES.

Program Benefits

As summarized in Table 4.1, the Elmira PEIP study found significant short- and long-term advantages for both the mothers and children in the intervention group. In the short-term, pregnancy behaviors were better for mothers in the intervention group, with reduced cigarette use, better nutrition, improved childbirth class attendance, and more social supports reported (Olds et al., 1986a). Intervention group mothers who did smoke bore 75 percent fewer preterm infants than did control mothers who smoked, and overall, intervention group teenage mothers bore heavier infants than the control group teenagers.

---

[7] See the Spring/Summer 1999 issue of *The Future of Children* (www.futureofchildren.org) for examples of other home visiting models, ranging from those that rely on lay professional home visitors to paraprofessional and professional home visitors.

[8] The Elmira model has been replicated by the same team of researchers in randomized trials in Memphis, Tennessee, and Denver, Colorado (Kitzman et al., 1997). The model is also being implemented at numerous other sites around the country.

The program assessment through age four showed that parental caregiving was affected by participation in the intervention. Reports of child abuse and neglect during the first two years of life were lower among the highest-risk intervention families (Olds et al., 1986b). Fewer safety hazards and more materials promoting development were found in the homes of the intervention group, and these children were seen less frequently in ERs through age four (Olds et al., 1986b, 1994.). Hospital days were significantly higher for the treatment group through age four, although this results from one outlier in the sample that appears unrelated to the program (Olds et al., 1994). Through age four, no significant differences in IQ, completed years of education for the mother, or home environment were found between treatment and control groups (Olds et al., 1986a, 1994).

The 15-year follow-up study found fewer reported acts of child abuse and neglect among the nurse-visited mothers for the full sample and the higher-risk sample (Olds et al., 1997). The other significant findings were restricted to the higher-risk sample (i.e., unmarried and low SES). For this group, months spent receiving AFDC and food stamps were significantly lower. The most at-risk mothers also had lower levels of criminal activity (measured by both self- and state-documented data on arrests, convictions, and jail days) and reported fewer behavioral impairments from alcohol and drugs. Although the treatment group also spent fewer months receiving Medicaid and more months employed, the differences were not statistically significant. The beneficial effects of the program in terms of controlling subsequent fertility continued through the 15-year follow-up, with treatment mothers reporting fewer subsequent pregnancies and births and a longer birth interval between the first and second child. Finally, children in the intervention group reported fewer arrests compared with the control group (Olds et al., 1997).

Cost-Benefit Analysis

A cost-benefit analysis of the Elmira PEIP was first undertaken by Olds et al. (1993) based on outcomes observed for participating children and their families through age four (i.e., about two years after the end of the intervention). Two years after the program ended, the analysis showed government savings that just exceeded program costs for low-income families (a net savings of $180 per child in 1980 dollars). For the sample as a whole, government savings did not exceed costs; rather, savings provided only a partial offset to costs. In both cases, the bulk of government savings resulted from reductions in the use of AFDC and other social welfare programs by the mother.

The most recent cost-benefit assessment was conducted by Karoly et al. (1998) based on the age 15 follow-up of program participants versus controls. Among the benefits for the PEIP, as shown in Table 4.1, only a subset were monetized for the cost-benefit analysis. They include savings to government from

- reductions in ER visits for the child through age four;
- reduced use of welfare by the mother through age 15 of the child;
- increased taxes from higher employment by the mother through age 15 of the child;
- reduced criminal justice system costs associated with the mother through age 15 of the child; and
- reduced criminal justice system costs for the child projected for the child's lifetime based on observed activity through age 15.

Benefits to the rest of society include the net increase in income associated with higher work effort by the mother (net of reductions in welfare payments) through age 15 of the child and the reductions in tangible crime costs associated with reduced criminal activity for the child projected over the child's lifetime based on observed data through age 15. As with the Perry Preschool cost-benefit analysis discussed above, all benefit streams were discounted to the birth of the focal child using a 4 percent real discount rate.

Karoly et al. (1998) compared the present discounted value of the government savings and benefits to the rest of society with program costs. As reported in Olds et al. (1993), the home visit program cost $3,246 in 1980 dollars. When converted to 1996 dollars to account for inflation, and when discounted to birth using a 4 percent real discount rate, the Elmira PEIP is estimated to have cost $6,083 in 1996 dollars per child served.

As noted above, the evaluation of the long-term follow-up results of the Elmira PEIP by Olds et al. (1997) focused on results for both the full experimental group as well as a higher-risk subsample consisting of unmarried mothers with low SES. In the results provided by Karoly et al. (1998), costs and benefits were analyzed separately for this higher-risk sample, as well as for the remaining experimental sample which was termed lower-risk.[9] The lower-risk group thus consists of two-parent or higher-SES families.

Table 4.3 summarizes the results of the Elmira PEIP cost-benefit analysis, with results reported separately for the higher-risk sample (top section) and the lower-risk sample (bottom section). Consider first the results for the higher-risk sample, which experienced the largest improvements in maternal and child outcomes as a result of participating in the program. The cost-benefit analysis indicates that the savings to government from changes in the mother's behavior and the child's behavior total $24,694, more than four times the program costs.[10] Another $6,072 in savings to the rest of society is generated in increased participant income and reductions in tangible crime losses. Overall, the net benefits of the program exceed $24,000, more than four times the program costs. About two-thirds of the more than $30,000 in total benefits is generated by savings to government from changes in the mother's behavior (largely a reduction in welfare costs), while the other third stems from changes in the child's behavior (primarily associated with reduced criminal activity). It is possible, as the children in the program make the transition to adulthood, that improvements in their economic outcomes (e.g., employment, welfare use) will generate additional savings that can be attributed to the child.

The results in the bottom section of Table 4.3 are not as encouraging for the lower-risk sample in the Elmira PEIP. For that group, the savings to government, based on those outcomes observed through age 15 of the child and that could be readily monetized, are less than

---

[9] As noted in Table 4.1, in the 15-year follow-up, the significant differences were primarily for the higher-risk families.

[10] As with the Perry Preschool program, the analysis of statistical uncertainty by Karoly et al. (1998) suggests that the net savings to government are positive for the higher-risk group but not for the lower-risk group.

Table 4.3

Costs and Benefits: Elmira PEIP

	Dollars per Child		
	Due to Mother	Due to Child	Total
Higher-Risk Families			
Program cost			6,083
Savings to government			24,694
Reduction in education services	*	*	
Reduction in health services	*	115	
Taxes from increased employment	5,683	*	
Reduction in welfare cost	14,067	*	
Reduction in criminal justice cost	634	4,195	
Additional monetary benefits			6,072
Increase in participant income net of welfare loss	1,010	*	
Reduction in tangible losses to crime victims	*	5,062	
Total benefits			30,766
Net benefits			24,683
Lower-Risk Families			
Program cost			6,083
Savings to government			3,775
Reduction in education services	*	*	
Reduction in health services	*	107	
Taxes from increased employment	1,144	*	
Reduction in welfare cost	1,270	*	
Reduction in criminal justice cost	111	1,143	
Additional monetary benefits			2,938
Increase in participant income net of welfare loss	1,622	*	
Reduction in tangible losses to crime victims	*	1,315	
Total benefits			6,713
Net benefits			630

SOURCE: Karoly et al. (1998), Tables 3.3, 3.4, and 3.7.

NOTE: * = not measured. All amounts are in 1996 dollars and are the NPV of amounts over time where future values are discounted to the birth of the participating child, using a 4 percent annual real discount rate.

$4,000 and are not enough to cover the program costs. The addition of nearly $3,000 in monetary benefits to the rest of society brings the total benefits to $6,713, just $600 more than the cost of the program. It is possible, however, that if other benefits of the program that are harder to monetize were included in the cost-benefit analysis, the net benefits would be even larger.

The cost-benefit analysis is not nearly as favorable for the lower-risk group because the program had a smaller impact in most of the domains captured in Table 4.3 compared with the higher-risk group (see Karoly et al., 1998, for additional detail). The lower-risk mothers and children, in many cases, had outcomes in the control group that were considerably better than their higher-risk counterparts, so there was less room for the program to change behavior. For example, in the absence of the PEIP, mothers in the lower-risk group spent 30 months on welfare in the first 15 years of the child's life, compared with 90 months for the higher-risk mothers. Although participation in the PEIP reduced welfare use even for the lower-risk mothers, the drop was to only 28 months. In contrast, higher-risk mothers in the program experienced an average of 60 months on welfare, a 30-month difference from the control group. This improvement generates $14,067 in savings to government for the higher-risk mothers compared with only $1,270 for the lower-risk mothers.[11]

THE CHICAGO CHILD PARENT CENTERS

The Chicago Child Parent Centers (CPC) program, a publicly funded school-based preschool and follow-on program, offers an interesting larger-scale contrast with the two model programs just highlighted (Reynolds, 2000). Operating continuously since 1967, the Chicago CPC initially provided a structured half-day program during the school year for three- and four-year-olds in 11 public schools in economically disadvantaged neighborhoods. In addition to preparing children for school through the promotion of reading and language skills, the program also provided comprehensive health and social services and promoted parental involvement. The program was expanded in 1978 to continue services through third grade, including a full-day kindergarten. Today, 24 centers provide preschool only or preschool and school-age components through grades one, two, or three.

---

[11] The savings in public assistance costs may not be as large in future replications of the PEIP because of the five-year lifetime limit that applies to receipt of public assistance for most adults under the welfare reform law passed in 1996.

In contrast to the two model programs discussed above, the evaluation of the CPC program is based on a quasiexperimental design with a group of 989 children who participated in the CPC preschool program for one or two years (and the CPC kindergarten) and a no-preschool comparison group of 550 children.[12] The treatment and comparison groups form a single age cohort that completed kindergarten in the spring of 1986. The latest follow-up took place in the spring of 2000 when the children were up to age 20 (Reynolds et al., 2000).

Program Benefits

Table 4.1 again summarizes the outcomes measured and results for the CPC program across the various follow-ups, with a primary focus on outcomes for the child. At the end of the intervention at age nine, those who participated in the CPC had significantly higher reading and math achievement scores, lower rates of grade retention, and higher ratings of parental involvement (1 = poor/not at all to 5 = excellent/much). No significant differences were found, on average, between participants and nonparticipants in special education placement and teachers ratings of school adjustment at age nine, although years of special education were significantly lower for treatment children by age 14 (Reynolds, 1994; Reynolds and Temple, 1995).

The differences in achievement scores between groups tended to become smaller over time, although they remained significant through age 14 for math scores.[13] Longer-term follow-up through age 20 revealed other lasting improvements, particularly in terms of educational outcomes (Reynolds et al., 2000). For example, years in special education by age 18 were lower for program participants, while rates of high school graduation and years of schooling completed by age 20 were higher.

---

[12] Some of the no-preschool comparison group eventually enrolled in the CPC school-age intervention. Thus, some results for the program are based on the sample of 1,150 children who participated in at least one year of the CPC program versus the 389 children who never participated in the program (Reynolds et al., 2000).

[13] The findings for regression-controlled mean differences are generally robust to those based on models that explicitly model selective program participation (Reynolds and Temple, 1995).

Researchers also examined measures of problem, illicit or illegal behavior in grades seven to 10, and again at age 20 (Reynolds, Chang, and Temple, 1997; Reynolds et al., 2000). Differences in delinquency rates between treatment and control groups and based on time in the program were significant at ages 13 to 14, but these were no longer evident at ages 15 to 16. However, by age 17, rates of petition to the juvenile court were significantly lower for participants.[14]

Cost-Benefit Analysis

Reynolds et al. (2000) have conducted a cost-benefit analysis for the Chicago CPC program based on data through the age 20 follow-up. Their analysis builds upon the methods adopted in Karoly et al. (1998) and Barnett (1993, 1996). All cost and benefit figures are expressed in 1998 dollars and discounted to age four of the focal child using a 3 percent real discount rate.

In particular, savings to government are calculated for the following outcomes observed for participating children:

- Reduced public education expenses due to lower rates of grade retention and reduced use of special education through age 18.
- Increased tax income projected from age 18 to 65 from greater earnings capacity due to higher rates of school completion at age 18.
- Reduced costs to the criminal justice system through age 17 of the child.

As with the Perry Preschool and Elmira programs, benefits to the rest of society were calculated in two domains: higher income for program participants projected through age 65 based on higher rates of high school completion through age 18 and reductions in tangible and intangible costs (e.g., pain and suffering) associated with lower levels of criminal activity observed until age 17.[15]

---

[14] Petitions capture criminal charges serious enough to be processed through the court system leading to possible sentencing by a judge (Reynolds et al., 2000).

The Chicago CPC program is estimated to have cost $9,931 per child for preschool plus follow-on services. This figure is based on an average annual cost of $4,520 for one year of preschool and $1,426 for one year of the follow-on program, including costs for personnel, equipment and supplies, capital expenditures, maintenance, and other outlays. About one-half of the participants enrolled in two years of the preschool program (for a cost of $6,933), while the average time in the follow-on program was about two years (for a cost of $2,998).

Table 4.4 reports the present discounted value of costs and benefits for the Chicago CPC program calculated by Reynolds et al. (2000). Similar to the results for the Perry Preschool program and the Elmira higher-risk sample, the Chicago CPC program generates total benefits nearly four times as great as program costs, a total of $36,613 in present discounted value benefits versus $9,931 in costs.[16] Savings to government alone are twice program costs, with most of the savings coming from lower education costs. The monetary benefits to the rest of society are driven by projected income gains for participants of nearly $12,000 (not accounting for any possible loss of welfare benefits).

---

[15] Note that, compared with the benefit calculations for the rest of society for the Perry Preschool and Elmira programs conducted by Karoly et al. (1998), the CPC calculations do not net out reductions in welfare benefits from the income gains to program participants. However, because reductions in welfare program costs are not counted as a benefit or savings to government, the net effect on total benefits to society is almost the same as would be calculated using the Karoly et al. methodology. The difference arises because Karoly et al. also account for savings in administrative costs in figuring the savings to government from reduced welfare program participation. In addition, in the CPC analysis, the crime savings include intangible benefits from reduced criminal activity, and the savings to government and the rest of society from reduced criminal activity are not projected beyond the observed age of 17. The CPC cost-benefit analysis also uses a lower discount rate (3 versus 4 percent) and discounts to age four of the child versus birth. Finally, dollar values are expressed in 1998 dollars rather than 1996 dollars. These differences mean that the results in Table 4.4 are not strictly comparable with those of Tables 4.2 and 4.3.

[16] Reynolds et al. (2000) did not report an estimate of the confidence interval for the net benefit result.

Table 4.4

Costs and Benefits: Chicago Child Parent Centers

	Dollars per Child		
	Due to Mother	Due to Child	Total
Program cost			9,931
Savings to government			19,970
Reduction in education services	*	10,585	
Reduction in health services	*	*	
Taxes from increased employment	*	3,300	
Reduction in welfare cost	*	*	
Reduction in criminal justice cost	*	6,085	
Additional monetary benefits			16,643
Increase in participant income	*	11,784	
Reduction in tangible losses to crime victims	*	4,859	
Total benefits			36,613
Net benefits			26,682

SOURCE: Reynolds et al. (2000), Figure 6.

NOTE: * = not measured. Results are for total CPC participation, which combines any preschool participation with any follow-on participation. Most of the cost savings result from the period of preschool participation. All amounts are in 1998 dollars and are the NPV of amounts over time where future values are discounted to age four of the participating child, using a 3 percent annual real discount rate.

LESSONS FOR FUTURE COST-BENEFIT ANALYSES OF EARLY CHILDHOOD PROGRAMS

Table 4.5 contrasts the results for the cost-benefit analyses of the three programs reviewed in this chapter. In particular, the table records the NPV of benefits minus costs for program participants, for the rest of society, and for the two groups combined, labeled *society as a whole.* All results are expressed in 1996 dollars to make them more comparable.[17] As discussed above, however, other differences in the cost-benefit methodology remain (e.g., discount rate, discount age, period covered by future projections), particularly for the Chicago CPC program generated by Reynolds et al. (2000) versus the Perry Preschool and Elmira PEIP results prepared by Karoly et al. (1998). Nevertheless, the comparison is instructive.

---

[17] The results for the Chicago CPC program were converted from 1998 dollars to 1996 dollars using the consumer price index (CPI-U).

All three programs demonstrate that the net benefits of early intervention can be sizable, especially when services are targeted to those who can benefit most. Net benefits to society exceed program costs by at least a factor of two, and upward of a factor of four. Program participants gain, especially when long-term follow-up reveals significant improvements in earnings for program participants compared with the control group (e.g., as in the case of the Perry Preschool age 27 follow-up and the Chicago CPC age 20 follow-up). These economic gains, projected for a full working career, are sizable even when discounted to the present. The benefits to the rest of society are also larger when early intervention programs lead to reduced lev-

Table 4.5

NPV of Benefits Minus Costs for Selected Early Childhood Intervention Programs

		Dollars per Child			
			NPV of Benefits Minus Costs for:		
Program (Cohort, N)	Dollars	Program Costs	Program Participants	Rest of Society	Society
High/Scope Perry Preschool (1962–1967, N=121)	1996	12,148	13,846	23,979	37,824
Elmira PEIP—higher-risk (1978–1982, N=100)	1996	6,083	1,010	23,673	24,683
Elmira PEIP—lower-risk (1978–1982, N=145)	1996	6,083	1,622	–993	630
Chicago CPCs (1967–present, N=1, 281)[a]	1996	9,559	11,343	14,340	25,683

SOURCE: Elmira PEIP and Perry Preschool: Karoly et al. (1998), Tables 3.3, 3.4, 3.7; Chicago CPC: Reynolds et al. (2000), Figure 6, with 1998 dollars converted to 1996 dollars using the Consumer Price Index (CPI-U).

[a] Results are for total CPC participation which combines any preschool participation with any follow-on participation. Most of the cost savings result from the period of preschool participation.

els of criminal activity in adolescence and young adulthood—improvements that can then be projected to continue into adulthood (e.g., as in the case of the Perry Preschool program and Elmira PEIP).

Each of the estimates reported in Table 4.5 are likely to be conservative for one reason or another. The intangible benefits for the rest of society from reduced crime levels are not included in the estimates for the Perry Preschool program or Elmira PEIP. Projected savings across adulthood from reduced criminal activity in adolescence are not included in the estimates for the Chicago CPC program. For all three programs, many of the benefits recorded in the evaluations have not been monetized (e.g., potential gains in health, changes in fertility behavior, and other life course changes as shown in Table 4.1). Finally, the evaluations also did not always measure outcomes in all the domains that might have been affected by the programs. For example, only the Elmira PEIP contained extensive measures of behavioral changes for participating mothers in such areas as education, labor market outcomes, welfare utilization, and criminal behavior. The Chicago CPC evaluation did not include measures of welfare utilization, while the Elmira PEIP assessments did not focus on educational outcomes for the child. Any potential benefits in these unmeasured domains would further add to the net benefits recorded in Table 4.5.

Other implications of the three cost-benefit analyses are discussed here for future analysis of other early intervention programs. Four issues in particular merit discussion.

Certain Outcomes Can Be Easily Monetized and Can Have Large Dollar Benefits. The cost-benefit analyses of the three programs reviewed here focused on a small set of outcomes that can readily be expressed in monetary terms and have the potential to generate large dollar benefits, either in terms of savings to government or for the rest of society. These include improved educational outcomes (e.g., as measured by special education use, grade repetition, school attainment), better labor market performance (e.g., as measured by work effort, earnings), reduced dependence on public assistance, and lower levels of criminal activity. Not all early intervention programs will significantly and substantially improve these outcomes for program participants—either children or parents—but those that do are likely to have a more favorable cost-benefit ratio.

Advantages of Long-Term Follow-Up. The three programs reviewed in this chapter provide among the longest follow-up periods for early intervention programs: at least to age 15 and up to age 27 for children who participated in intervention programs starting as early as birth. Most important, long-term follow-up allows assessment of program impacts in domains that can be readily monetized, such as those identified above: educational performance, labor market success, public assistance utilization, and criminal activity. These outcomes are not observed for participating children immediately after an early intervention program ends. Instead, participants (and controls) must be followed into adolescence and beyond to capture benefits in these domains. Many of the outcomes observed for children during the period of program delivery and shortly after an early intervention program ends are in such areas as cognitive and behavioral functioning, which are not easily translated into dollar benefits for participants or the rest of society (see Table 4.1).

One disadvantage of long-term follow-up is that conditions may change considerably between a program's implementation and when the long-term effects are known. The evidence that the Perry Preschool program was a good societal investment in the early 1960s is strong circumstantial evidence but not proof that a replication today would also be a good investment. Much has changed in the intervening four decades.

Some Benefits Can Be Projected Beyond the Period of Follow-Up. In some cases, we have a good understanding of how outcomes at younger ages are related to outcomes at older ages. For example, based on criminal activity observed through adolescence, it is possible to predict the future profile of criminal behavior through adulthood. Likewise, earnings and public assistance utilization trajectories in young adulthood can be used to forecast experiences during the entire work life. Educational attainment can also be used to project lifetime earnings profiles. Thus, with longer-term follow-up, benefits observed through the age of follow-up can be projected further into the future. These added benefits, even when discounted to the present, raise the benefit-cost ratio for an early intervention program. These projections, however, introduce additional uncertainty into cost-benefit analyses and are not as readily supported in other outcome domains.

Changes in Parental Behavior May Generate Benefits Soon After a Program Ends. While longer-term follow-up is required to observe changes in behavior in relevant domains for participating children, benefits from potential changes in parental behavior may be realized when children are younger. For example, the Elmira PEIP, which was designed to affect the life course of participating mothers, produced improvements in their outcomes in such areas as labor market activity, public assistance utilization, and criminal behavior. Karoly et al. (1998) show that the cumulative present discounted value of savings to government for the Elmira higher-risk sample actually exceeds program costs by age three of the child, just one year after the program ended. This "break-even point" is reached so rapidly because of immediate changes in the mother's behavior that generate sizable savings. In contrast, Karoly et al. (1998) calculate that the break-even point for the Perry Preschool program is not reached until about age 20 because savings to government are calculated only for changes in the child's behavior in domains not realized until adolescence and young adulthood.[18] It is possible that the Perry Preschool program would have an earlier break-even point if savings from improvements in parent's outcomes could be measured and incorporated into the cost-benefit analysis.

---

[18] The Chicago CPC cost-benefit analysis by Reynolds et al. (2000) does not include a calculation of the break-even point.

Chapter Five

APPLYING COST AND OUTCOME ANALYSIS TO THE *STARTING EARLY STARTING SMART* PROGRAM

This chapter applies the methods outlined in the previous chapters to the *Starting Early Starting Smart* (SESS) program. We consider both data now being collected by SESS and potential options for future data collection and program design. This exercise not only informs SESS policymakers about the use of current data and future opportunities for analysis, but it also helps illustrate how the methods discussed can be put into place for a real-world programs.

We begin this chapter by describing the SESS program. Then we outline approaches to analyzing cost and outcome data for the program. We also discuss some key methodological considerations relevant to conducting cost and outcome analysis for this program.

THE SESS PROGRAM AND EVALUATION DESIGN

SESS is designed to test the effectiveness of integrating behavioral health services for children from birth to age seven and their families, relative to the outcomes for children and families who receive the usual standard of community care. Integrated behavioral health services are defined as substance abuse treatment, substance abuse prevention, and mental health services.[1] The initial four-year phase of the SESS program—Phase I—began in 1997.

---

[1] This discussion of the SESS program and evaluation design draws on the *Starting Early Starting Smart Phase One Report*, prepared by the SESS Data Coordinating Center, August 1998.

SESS currently has cooperative agreement grantees in 12 sites nationally. These sites fall into two natural clusters based on their organizational settings—primary health care (PC) and early childhood development (EC). PC sites provide health care to families of target (index) children, and EC sites provide preschool education services to index children. There are currently five PC sites and seven EC sites. (See Appendix A for a full list of SESS sites and a brief description of their program features.) These clusters vary in several important ways, as shown in Table 5.1. PC sites specifically target moderate- to high-risk families. However, participants at EC sites also generally demonstrate relatively high levels of stress and risk factors.

SESS is purposefully designed as a multisite study encompassing diverse field settings in hopes of generating strong evidence of its general applicability. In addition to units of observation at the program level (PC and EC), the units of analysis for the individual level are the index child and the family. The logic behind the design is twofold:

Table 5.1

Characteristics of SESS PC and EC Demonstration Sites

Primary Care	Early Childhood
Intervention begins from birth to age 3 in most sites	Intervention begins from ages 3 to 5 in most sites
Eligibility is based on individual screening to target caregivers or children who have specific risk behaviors	Eligibility is based on the setting, not the individuals within it; entire classrooms are eligible for these services (e.g., Head Start)
Program focuses behavioral health resources on parent	Program focuses on behavioral health and developmental needs of index child
Needs of caregiver determine program participation	Needs of caregivers are evaluated more indirectly
Case management component is an innovative addition in this setting	Behavioral health component is an innovative addition in this setting
Experimental design is used for all sites	Quasiexperimental design is used for all but one site, which is experimental

- Use an experimental or quasiexperimental design to detect program effects at the individual level, and
- Use variation in target population, program context, or program intervention at the program level to explain differences in program effectiveness across sites.

The sample sizes vary across sites, but most are around 100 to 300 index children. The pooled sample consists of 1,584 persons in the treatment group and 1,303 persons in the control (or comparison) group.

The current SESS evaluation is designed to test two specific hypotheses:

- The integration of behavioral health services within PC or EC service sites will lead to higher rates of entry into prevention, early intervention, or the treatment of children/families identified as in need of services (also greater participant satisfaction).
- The integration of behavioral health services within PC or EC service sites will lead to improvements in social, emotional, and cognitive functioning in children and families served.

The first hypothesis focuses on outcomes of services access and utilization and satisfaction, while the second focuses on family functioning, parent-child interaction, and child outcomes.

SAMHSA and CFP have funded a set of cross-site data activities that include data collection, manipulation, and analysis. As part of these activities, they have mandated the creation of an overall program database. The five types of data collected as part of this database include site-level intervention descriptions, contact log data (collected only for the treatment group), Services Access and Utilization and Satisfaction Survey, baseline data, and outcome data. These measures are collected at baseline and for an 18-month follow-up period, with follow-up intervals that average six months (PC sites) or nine months (EC sites). Baseline data and some follow-up data have been collected for treatment and comparison groups. While most sites have attempted to include a comparison group, some sites include no comparison group or a comparison group that receives some SESS services.[2] CFP and SAMHSA are considering funding a longer-term follow-up for participants in a subset of the current sites. Currently, no cost data are being collected in Phase I, nor are the SESS evaluation design and the longer-term follow-up currently incorporating cost-benefit or related analysis.

CFP and SAMHSA plan to implement a second phase of the SESS program (Phase II), which is currently being designed. Assessing the feasibility of including cost and outcome analysis is part of the planning process for Phase II. In the remainder of this chapter, we assess the utility of data being collected in Phase I for this type of analysis and make recommendations for alterations to the Phase I design, which could be implemented in Phase II.

USING THE SCORECARD AS A FRAMEWORK

As a framework for our discussion of potential cost and outcome analyses for the SESS program, we return to the scorecard introduced in Chapter Two of this report (Table 2.1). By characterizing the cells of the scorecard that can be filled in with Phase I data, we can assess the types of analysis that could be conducted with the data currently being collected, and we can identify additional data that would need to be collected in the next phase.[3]

As discussed in Chapter Two, a number of types of cost and outcome analyses could be undertaken for such a program as SESS. Specifically, at least three broad types of analysis could be conducted for this program:

- Cost-savings or cost-benefit analysis, whereby the costs of the program are compared to the benefits of the program from the perspective of the government and society at large, respectively.
- A type of cost-effectiveness analysis, which compares the change effected by different variants of the PC sites or the EC sites or

---

[2] See Appendix A for more information about each site's comparison group.

[3] In making our recommendations, we do not explicitly discuss a number of the methodological issues described in Chapter Two, such as choice of discount rate and accounting for statistical and scenario uncertainty. These can be addressed during the cost analysis, following standards established in the cost-benefit literature.

examines which design features of SESS programs were associated with the greatest "bang for the buck."

- Characterization of the costs of implementing SESS so that future sites hoping to replicate the program have reasonable expectations regarding the costs they would incur.

While other approaches could certainly be enumerated, these three represent the general classes of analysis best aligned with the stated objectives of the policymakers for this program.[4]

As we proceed in the remainder of this chapter, we rely on the scorecard framework to make a series of recommendations about the evaluation design and the collection and analysis of cost and outcome data. However, a number of our recommendations specific to cost and outcome data depend, in part, on the type of analysis desired for the SESS program. This in turn will reflect the objective that the analysis is trying to achieve, such as the three listed above. For example, if the goal of the cost and outcome analysis is to characterize the costs of implementing SESS for potential future replication, the bulk of the cost data would pertain to the costs to the agency implementing the program. However, if the goal is a comparison of the costs and benefits of the program from the perspective of society at large, then a more comprehensive enumeration of the costs and outcomes of the program would be required. We revisit these issues again at the end of this chapter.

> **Recommendation** : Specify the explicit goals of the cost and outcomes analysis to guide the scope of cost and benefit data collection and analysis.

Defining the Baseline and Alternative Policies

We first need to establish the columns of the scorecard—i.e., what would serve as the baseline comparison group and what would serve as the alternative programs. As discussed above, the baseline represents the world without the SESS program elements.[5] In the case of the SESS Phase I design, there is a baseline case associated with the two basic program models: primary care (PC) sites without SESS and early childhood (EC) sites without SESS. An SESS information packet states that grantees are required "to address the multiple needs of poor and at-risk families and their very young children by providing coordinated, wraparound services, with special emphasis on services that address the participants' behavioral health needs." Hence, the marginal contribution of SESS is the integrated mental health and substance abuse prevention and treatment services delivered in these settings, plus coordination activities that may change the amounts of other services that participants receive. SESS's marginal contribution is not the entire range of services provided at these sites. This is why the comparison group is PC sites or EC sites without SESS rather than a control group that receives no services of any type, including PC or EC services.[6]

---

[4] These objectives and other issues related to the application of cost and outcome analysis to SESS are described in Cannon, Karoly, and Kilburn (2000). This document summarizes a meeting held between SESS funders and program staff and experts in cost analysis from both RAND and other organizations.

The alternative programs under consideration are the PC and EC sites with SESS. However, the Phase I demonstration of SESS was purposefully designed to have variation within the PC and EC models in the treatment populations and suite of services offered to participants across the demonstration sites. As a result, there is a baseline for each combination of geographic site and program model. Thus, it would be possible to consider a number of variations of SESS PC and EC sites to assess how differences in the population served and/or the services provided influenced costs and outcomes. This corresponds to the second type of cost and outcome analysis—cost-effectiveness analysis—enumerated above.

For the sake of brevity, in the remainder of this discussion we will assume that for our hypothetical example there is only one variant of

---

[5] It is also possible to design an evaluation with a baseline that represents a world with no program at all, either the basic services offered at PC or EC sites or any of the add-on elements of the SESS program. In this case, the costs and benefits of both basic PC or EC services plus the SESS overlay would be compared with a control group that received no SESS, EC, or PC services.

[6] In the Phase I implementation of SESS, those in the control or comparison group at the PC sites receive services from the same PC provider that also offers integrated SESS services to the treatment group. It is possible that even the basic PC services are changed as a result of the provider offering the integrated SESS services for the treatment group, for example, stemming from the capacity building of the staff, and so on.

PC with SESS (PC plus SESS) located in one geographic site, but two variants of EC with SESS in two separate geographic sites, which we shall call EC1 plus SESS1 and EC2 plus SESS2. In Table 5.2, we show how the columns in the scorecard would appear for this set of comparisons.

The consideration of comparison groups and policy alternatives raises four design issues for the planned Phase II evaluation of SESS. The first is the use of an experimental versus quasiexperimental design, i.e., whether the baseline is a randomly assigned control group or a matched comparison group. The Phase I design (see Table 5.1) includes a mix of sites, some with random assignment (primarily PC sites) and others with matched comparison groups (mostly EC sites). Preliminary data from the evaluation raise concerns about the preintervention comparability of the matched comparison groups in the EC sites (see the summary of the discussion in Cannon, Karoly, and Kilburn, 2000). If such differences exist, any postintervention differences between the treatment and comparison groups may be due to other factors besides the SESS services. To obtain the best research results, random assignment would be used for the evaluation design at all sites in a subsequent demonstration phase, if at all possible.

However, random assignment may not be feasible for several reasons. As we pointed out earlier, results of early childhood interventions can be extremely sensitive to the risk characteristics of the population they serve. They may have big effects when applied to high-risk children, but smaller effects when applied to lower-risk

Table 5.2

Illustrative Scorecard for Hypothetical SESS Example: Alternative Policies

	Alternative Policies					
Impacts	PC Only	PC+ SESS	EC1 Only	EC1+ SESS1	EC2 Only	EC2+ SESS2
Program Descriptors						
Cost Elements						
Outcomes						

children. Random assignment means refusing program services to some high-risk children, and this may be difficult to do in certain settings. In the case of the SESS program, this may be more of a concern for the PC sites where treatment and control children are served by the same provider. At EC sites, this may be less of a concern, since the SESS services are offered to whole classrooms of children rather than to randomly selected individuals. Likewise, control groups consist of whole classrooms to which SESS services are not offered. One can reasonably expect to find children at all risk levels in both the control and participant classrooms. But at both kinds of sites, if random assignment is not possible, it is important to match controls to participants in terms of risk factors.

> **Recommendation** : Where possible, use random assignment to define control groups in order to provide a more valid test of SESS program effects. When random assignment is not possible, strive to match children in the treatment and comparison groups in terms of their risk factors.

A second issue concerns data collection for the control group. In Phase I, participants and controls alike received an initial interview and several follow-up interviews at intervals that average six to nine months for PC and EC sites, respectively. For each participant, however, each SESS site keeps a contact log that describes every telephone contact and every face-to-face contact with SESS staff. Data this complete and detailed are not available for controls. In particular, it is not known, save by self-report after delays of several months, just what services the controls are receiving. They may, in fact, be receiving many of the same services as the participants. It might be possible to obtain more complete and accurate records of services received by controls from records kept by the service providers. Of course, controls would have to provide consents for SESS to gain access to these records.

> **Recommendation** : Strive to collect service, cost, and outcome data on the control groups that are as complete as the data on the treatment groups.

A third issue concerns the extent of variation in the SESS program models as implemented across demonstration sites, both in terms of the services provided and the target population served. In the Phase

I design, the program models, and to some extent the population served, vary by geographic site even within the PC and EC program models. This variation can be useful for identifying the most successful program designs based on the Phase I outcomes data. However, it is difficult to disentangle differences in program effectiveness stemming from the program model, geographic site, or population served. For Phase II, there are advantages to considering a more limited set of the best designs that emerge from Phase I, possibly implementing the same program model in two geographic sites or for different target populations or implementing two different models in the same geographic site or for the same target population. Alternatively, it may be desirable to fix the target population, selecting among the at-risk groups identified in Phase I that benefit the most from the SESS program model. In either case, for an evaluation of a given total sample size, a more refined and uniform program model in Phase II will allow the evaluation to consider how outcomes and costs vary with the characteristics of the site, target population served, or program model. This will be important information to guide future program implementation.

> **Recommendation** : In Phase II, impose more uniformity in the program models across sites, strategically selecting a few variations in design based on outcomes data from Phase I.

A fourth important consideration that influences the viability of conducting cost and outcome analysis for SESS is the ability of the Phase I or planned Phase II evaluations to retain subjects (both control and treatment group members) across time. This is important because attrition from evaluation studies is rarely random. Instead, those who continue to receive program services or to be assessed in terms of their outcomes are likely to differ from those who drop out of the program or are lost to follow-up in ways that may not be controlled for by differences observable to the researcher. Analyzing data that contain only the individuals who remain in the program over time and who continue to be monitored could generate misleading conclusions regarding the effectiveness of the program. In the first follow-up of Phase I data collection, participant retention from the initial survey ranged from nearly 99 percent to a low of 56 percent across sites, with mean retention in the EC sites and PC sites of 82 and 61 percent, respectively.[7] Because of the importance of collecting long-term outcomes for children's intervention programs, this issue also merits special attention during the Phase II design.

> **Recommendation** : Use the information from the Phase I evaluation to assess the reasons for attrition from the study. In Phase II, devote more resources to retaining study subjects, remedying the retention problems identified for some sites in Phase I.

Describing SESS Sites

Now we turn to filling in the rows that should be described under the three broad headings in Table 5.2. The first information we need to specify are the features, or "program descriptors," of each baseline program and each alternative policy. They should be detailed enough so that future sites, which may be considering implementing variations of the policies, could have a reasonable expectation of replicating the conditions under which the costs and outcomes were realized.

While a complete list of program descriptors may include dozens of entries or more, we list types of information here that would be candidates for inclusion:

- Population served, especially including risk category or characteristics that determine risk. Eligibility criteria should be listed as well.
- Characteristics of personnel providing services (such as education, certification, and bilingual skills).
- Typical services received by participants (such as a particular substance abuse prevention curriculum, enriched preschool that focused on specific skills, psychiatric evaluation, medication monitoring, and residential substance abuse treatment).
- "Dosage" of services, including number of visits and length of visits of various types. Note that services provided will generally

---

[7] Documentation provided by the SESS Data Coordinating Center based on response rates as of December 12, 2000.

be tailored to the population served, so types and dosages of services will need to be specified separately for different population subsets.

We indicate some illustrative program descriptors for our hypothetical SESS example in Table 5.3. Note that ideally, the features of the "baseline" or comparison program should be as close as possible to those of the "treatment" program, save for the specific features that characterize the SESS program.

When characterizing the program features, it is important that they be based on information on how a program is actually implemented, not just on the planned design. In the Phase I evaluation of SESS, a component of the data collection includes site visits to gather information about how each program model is actually operating. This is critical information required for conducting a valid comparison across program models and should be continued in the Phase II

Table 5.3

Illustrative Scorecard for Hypothetical SESS Example: Program Descriptors

	Alternative Policies					
Impacts	PC Only	PC+ SESS	EC1 Only	EC1+ SESS1	EC2 Only	EC2+ SESS2
Program Descriptors						
Population served						
Age of child at enrollment in program						
Eligibility criteria						
Transportation provisions						
Child care available during parent appointments						
Health services provided						
Etc.						
Cost Elements						
Outcomes						

design. This information is also useful for ensuring fidelity to a program model as designed, so that "program drift" is minimized and "dosage" levels are maintained.

> **Recommendation** : In Phase II, continue to collect information on program features through site visits and other mechanisms to characterize accurately features of the intervention models as they are implemented and to ensure fidelity to the program model.

COLLECTING AND ANALYZING SESS PROGRAM COSTS

The second broad heading shown in Table 5.2 is cost elements. The cost of the SESS program would entail a comparison, for each program model, of the costs with and without the SESS component. That is, the costs of the PC plus SESS programs would be the difference between the costs of the PC model without SESS and the costs of the primary care model with SESS. Similarly, the costs of the EC plus SESS programs would be the difference between the costs of those programs with and without the SESS component. This comparison thus requires collecting cost data for both treatment and control group participants at each site where SESS is implemented. Collecting cost detail at the level of each participant is possible, but this can be time-consuming. It is probably sufficient, for most analyses of SESS that would be of interest, to construct aggregate program costs at each site, rather than cost disaggregated by participants or groups of participants at each site. The most likely exception would be if high-risk children were provided much more intensive services, or were retained longer in the program, than low-risk children.

> **Recommendation** : Collect cost information for both treatment and control groups at each site where SESS is implemented.

The cost principles outlined in Chapters Two and Three should guide the completion of this section of the scorecard. In particular, information characterizing the following categories should be enumerated in the scorecard:

- Resource categories. These include personnel, equipment, facilities, and supplies/other.
- Explicit expenses and in-kind costs.

- Fixed and variable costs.
- Consumable and nonconsumable items.
- Investment costs and operating costs.
- Stakeholder group. Such as participants, the agency implementing the program, or society at large.

Rather than including a row for each combination of these various categories, a good start would be to include sections for stakeholders and resource categories. As discussed earlier, the following groups are likely to incur costs as a result of the program:

- Participants. Their costs may include time and resources getting to appointments, child care while the parents are in meetings or appointments, the value of the time spent in appointments, and others.
- The Agency Implementing the Program. The agencies' costs will include the labor bill for staff, the rent or space costs, such operating costs as utilities, supplies and equipment, and others.
- Other Agencies or Providers. These may include public or private agencies or providers to whom SESS participants are referred for services, such as special education services or family violence prevention programs.
- Society as a Whole. The costs to other components of society might be the value of the time of volunteers at the agency implementing SESS, donated space or supplies, or the value of the public infrastructure, such as public transportation, which may play a role in the delivery of SESS services.

We have shown these four groups of stakeholders, which might accrue costs, in italics in Table 5.4. As noted in Chapter Three, it is critical that identical cost information be collected for both treatment and control groups for each of the parties listed above. This allows investigation of possible cost-shifting or cost-offsets that otherwise might go undetected.

Also in this table, we have listed a few examples of resource categories for the two groups of stakeholders—participants and agencies implementing the program—as an illustration. We have also included a couple of examples of specific items, which might be included in the rows. Since participants are unlikely to incur facilities costs or equipment costs as a result of participating in SESS, we have only included personnel and supplies/other categories for participants. A much richer list of cost entries would need to be developed for each stakeholder and each resource category as part of the analysis of the SESS program. Once the particular items that go in the rows have been identified, they can be demarcated according to the other characterizations enumerated above, such as explicit expense or in-kind expense, investment cost versus operating cost, and so on.

> **Recommendation** : The cost information should be as comprehensive as possible. Costs borne by various parties by resource category should be differentiated; the time period that costs are incurred should be identified; and direct and indirect costs, fixed and variable costs, and goods and services provided in-kind should be measured.

Currently, SESS data collection efforts in Phase I focus on outcome measurement and do not include data on costs. Even though such issues as the quality of comparison groups are not likely to be resolved in Phase I, collecting cost information for the extension sites in Phase I would still have great utility, particularly for informing the Phase II design. For instance, if different types of PC plus SESS or EC plus SESS sites realized similar outcomes, but one type of either PC or EC site had half the costs of the others, policymakers may want to focus Phase II investments in the lower-cost option. Similarly, collecting data in the Phase I extension sites might help identify specific program features that have the greatest impact on key outcomes in relation to cost per family served. Again, this could help suggest which program features Phase II should emphasize or encourage. Beginning to collect cost data for the Phase I extension sites would have the additional advantage of serving to work out data collection procedures before Phase II, and to indicate how much of the Phase II evaluation budget should be set aside for the collection and analysis of cost data.

> **Recommendation** : Collect cost data for the Phase I extension sites to inform the design of Phase II and help prepare for Phase II cost data collection.

Table 5.4

Illustrative Scorecard for Hypothetical SESS Example: Cost Elements

	Alternative Policies					
Impacts	PC Only	PC+ SESS	EC1 Only	EC1+ SESS1	EC2 Only	EC2+ SESS2
Program Descriptors						
Cost Elements						
Participants						
Personnel costs						
Lost work time						
Supplies/other						
Transportation resources						
Child care costs						
Agency Implementing Program						
Personnel costs						
Number of hours spent per nurse home visit, including preparation, travel, follow-up, etc.						
Number of nurse home visits per participant						
Equipment						
Computer and related equipment						
Facilities						
Rent						
Utilities used						
Supplies/other						
Travel costs for nurse home visit						
Other agencies or providers						
Society as a whole						
Outcomes						

Finally, in collecting cost information, whether for Phase I extension sites or Phase II sites, it is important that the data collection procedures be as uniform as possible across SESS demonstration sites, with all sites capturing costs for the same parties, cost elements, and time periods. This is implicit in the construction of the scorecard, yet it is still worth emphasizing given that the capacity for data collection and the cost accounting systems may be quite different across sites. A critical element in the collection of cost data will be appropriate training and support at each site and for any data collection organization that may operate across sites. The cost associated with training for and gathering cost information (and the outcome information discussed below) should also be collected. If data collection becomes a standard part of implementing the SESS model, this information will allow these costs to be incorporated into the estimate of the full program costs. Alternatively, if future implementation of SESS will not require detailed data collection, or only a more streamlined data collection procedure, the program costs can be adjusted accordingly. The same is true for the cost associated with the analysis of the cost and outcome data collected.

> **Recommendation** : Plan for proper training and technical support of SESS sites and any cross-site data collection organizations to ensure uniformity in the collection of cost data. Collect information on the cost of data collection, training and support, and the related analyses of the data.

COLLECTING AND ANALYZING SESS PROGRAM BENEFITS

The final heading shown in Table 5.2 is program outcomes. Like the cost elements in the scorecard, the outcomes in the scorecard would also need to demarcate the individuals to whom benefits accrue and the period when gains are realized. The benefits of early childhood intervention programs have typically been measured for program participants in the four broad domains reviewed in Chapter Three: emotional and cognitive development, education, economic well-being (e.g., public assistance receipt, income, crime), and health. The specific outcome measures in each category—and whether they are measured for participating children, parents, or both—is a function of the program design and the expected areas of impact. As noted in the discussion in Chapter Three, some of these impacts—such as those in the economic sphere and a subset of those in the education domain—when applied to children require longer-term follow-up to observe changes in their outcomes at more advanced ages, long after the intervention has ended.

Chapter Three also highlighted some impacts that result from changes in participants' behaviors that can also affect outcomes for nonparticipants. For example, reduced criminal activity on the part of participating parents or children produces benefits to other members of society in the form of lower crime rates. Another example: Improved behavior of program participants during their school-age years may improve classroom learning for other children at school. Likewise, improved outcomes for the parent may have spillover benefits for the parent's other children in addition to the target child in the intervention.

The current data collection effort for the first phase of SESS is guided by the expected areas of program impact and an evaluation initially planned based on a two-year period of data collection. In particular, the SESS evaluation focuses on multiple domains of expected impact: access, utilization, and satisfaction with behavioral health services and family functioning; parent-child interactions; and child outcomes. Data currently being collected include measures of the following:

- Focal child characteristics.
- Family/household characteristics.
- Parent/caregiver characteristics, such as demographics, education, employment, public assistance, insurance, etc.
- Child problem behavior and social skills.
- Child cognitive development.
- Parent-child interaction.
- Parent/caregiver stress and negative/positive behaviors.
- Parent/caregiver mental health problems.
- Home environment, such as safety/violence and learning opportunities.
- Service utilization and satisfaction.

As indicated by this list, the SESS evaluators are collecting outcome data for both parents and children.

For purposes of the various cost and outcome analyses, the outcomes being collected for the SESS evaluation do not include most of the measures italicized in Table 3.2, i.e., those most readily translated into monetary benefits, either to government (taxpayers) or to other members of society. In fact, many of the above outcomes—which largely fall in the class of cognitive or emotional development measures—would be difficult to translate into monetary terms. Other benefits, such as better access to needed services or more appropriate use of health care services, are also difficult to express in monetary terms. This makes a formal cost-benefit or cost-savings analysis problematic in that only a limited set of outcomes might possibly be valued in dollar terms to be compared with program costs. Unless the program impact for those outcomes that are monetized is very large and favorable, so that sizable dollar benefits are generated, it is unlikely that a cost-benefit analysis would show a favorable outcome for the SESS program based on the information available after two years.

Given the current data collection plan for Phase I, cost-effectiveness analysis, which compares the change in outcomes elicited by a program to the costs of the program, is feasible provided cost data are assembled for the current or extension sites. This is because the outcomes are not translated into dollar terms but rather remain in their natural units, such as values on a given scale. Because no summary cost-benefit measure is generated, however, this approach requires decisionmakers to weight the various outcomes using their own subjective weights. Another type of analysis, which could be executed with the currently available data, is an assessment of which design features of programs yielded the greatest influence on outcomes. This type of analysis is currently planned as part of the Phase I evaluation.

> **Recommendation** : If Phase I cost information can be collected as recommended above for Phase I extension sites, focus cost and outcome analysis based on Phase I data on cost-effectiveness measurement.

If the objective of the cost and outcome analysis is to perform cost-savings or cost-benefit analysis, it will be important to broaden the types of short-term measures collected, especially for parents and other caregivers, and to consider an evaluation with a longer-term follow-up. As demonstrated by the cost-benefit analyses in Chapter Four, parents' outcomes have the potential to produce the largest short-term gains as the result of an early childhood intervention program. In contrast, improvements in children's outcomes may take years or even decades to reveal themselves. For this reason, if analysis that compares the benefits and costs of SESS is desired, collecting longer-term outcomes in Phase II would be valuable. While modeling is able to predict some longer-term outcomes based on observed changes in outcomes in the short run, obtaining data over the longest period possible avoids the statistical uncertainty inherent in such forecast modeling. The scenario uncertainty remains, of course.

A possible longer-term follow-up of the Phase I or planned Phase II demonstration sites would allow for a broader set of measures to be collected for participating children and their parents, including those that might produce larger impacts or impacts that can at least be monetized. The cost-benefit analyses of the early childhood programs reviewed in Chapter Four demonstrate the value of collecting information in the short- and medium-term (e.g., two to 10 years) for parents and in the longer-term (e.g., 10 to 20 years) for children on outcomes such as public assistance program use, employment, earnings, and criminal activity. If behavioral changes are large in these areas as a result of the SESS intervention, they can produce sizable dollar benefits that, even when discounted, will be a large offset to the costs of the program.

Table 5.5 illustrates some of the outcome measures that might be used for longer-term follow-up of the SESS program. The key outcome areas discussed in Chapter Three that are easily expressed in dollar terms are represented, and measures for both children and adults are assessed as of a specific age, *A*, of the focal child. Whether or not the SESS program will produce outcome gains in these areas has yet to be determined, but there is reason to believe that increasing access to substance abuse treatment services and mental health services will affect at least some of these domains. Substance abuse has been found to impose huge economic costs on society (Rice et al., 1990), and treatment has been demonstrated to be more effective than either no treatment or incarceration (McLellan et al., 1996). Other research has found that over 90 percent of the total cost savings produced by substance abuse treatment is in the form of reduced criminal justice system costs (see, e.g., CSAT, 1999). Moreover, in a comparison of treatment to other cocaine control programs, Caulkins et al. (1999) showed that treatment was more cost-

Table 5.5

Illustrative Scorecard for Hypothetical SESS Example: Outcomes

	Alternative Policies					
Impacts	PC Only	PC+ SESS	EC1 Only	EC1+ SESS1	EC2 Only	EC2+ SESS2
Program descriptors						
Cost elements						
Outcomes						
Outcomes for children						
Number of grades repeated through age *A*						
Years of special education through age *A*						
Years of education through age *A*						
Months employed through age *A*						
Average monthly earnings at age *A*						
Months receiving public assistance through age *A*						
Number of arrests through age *A*						
Emergency room visits through age *A*						
Etc.						
Outcomes for parents						
Years of education through age *A*						
Months employed through age *A*						
Average earnings at age *A*						
Months receiving public assistance through age *A*						
Number of arrests through age *A*						
Etc.						

effective than other approaches, including prevention, enforcement, and interdiction.

Improvements in mental illness rates would be expected to yield gains in labor force outcomes given that the percentage of persons out of the labor force and unemployment rates are significantly higher for persons with mental disorders (Sturm et al., 1999). The most comprehensive evidence on mental health services that explicitly incorporates cost-outcome methods is for the assertive community treatment (ACT) program, which provides services for those with serious mental disorders. Results indicated that subjects in the experimental group had improved outcomes compared to the control group and that family and community burden did not increase. Given increased wages and lower income support for the experimental group, societal costs were found to be slightly lower than for the control group (Test and Stein, 1980; Stein and Test, 1980).

These findings were countered by results that showed the results of the two groups converged after the program was terminated (see discussion in Hargreaves et al., 1998). Given the focus of the SESS intervention on increased access to and utilization of substance abuse and mental health treatment services, the SESS program could also produce benefits in similar areas.

It may also be fruitful to collect information in other outcome domains for possible inclusion in a cost-benefit analysis. For example, information on educational outcomes for children may be collected as early as the primary grades, with possible improvements in such outcomes as lost school days, grade repetition, and special education use that can be valued and tallied against program costs.

For parents and other caregivers, improvements in physical and mental health or reductions in such outcomes as family violence and child abuse and neglect may be evident in the short and medium term. These outcomes can potentially be valued as well in terms of increased labor market productivity and reduced use of other health care services. Again, it is not certain that the SESS program will significantly affect these outcomes, but they are among the likely candidates for improvement, and they can be translated into monetary benefits for participants or other members of society. Given the opportunity costs associated with added data collection, any new measures collected should be selected based on a theoretical model of the SESS program's expected impacts along with evidence that similar interventions have produced gains in those areas.

> **Recommendation** : If cost-benefit or cost-savings analysis is the objective for SESS, then outcome data should be supplemented to include information for parents and other caregivers in the short and medium term in the domains of health and economic well-being (e.g., labor market outcomes, public assistance use, criminal activity, and justice system contact) and for children in the medium term in the domain of educational outcomes and longer term also in the domain of economic well-being. The choice of specific outcome measures should be guided by findings from related evaluation studies whenever possible.

If a longer-term evaluation study is designed or anticipated for either Phase I or Phase II, several methodological issues discussed in Chapter Three should be considered. First, if a long-term follow-up is anticipated at the outset of the evaluation, it is important to collect information that will ensure the lowest possible rates of attrition and that allow data collection through administrative sources along with survey data. This would include, for example, obtaining identifying information for program participants, such as Social Security number or driver's license number, at the outset of the intervention. This would allow tracking of those in the treatment and control groups for subsequent follow-up interviews or searches for data in administrative databases (e.g., employment histories, criminal records).

> **Recommendation** : For a Phase I follow-up or Phase II design, obtain information from participants that allows collection of administrative data and permits effective tracking of individuals to increase response rates at later follow-ups.

Second, as discussed in Chapter Three, it is desirable to collect complete histories for some outcomes that may generate a continuous flow of dollar benefits. Thus, for example, if employment outcomes are better each year after an intervention ends, it would be ideal to know about employment rates in each year since the last follow-up in addition to their current status. A complete history of public assistance program use or use of costly special services in education or health care would also be relevant. Depending on the interval since the last follow-up, it may be difficult for respondents to recall a complete history, but such retrospective information can be of high quality when the events recorded are particularly salient. Administrative data, when available, also often provide a complete history with less concern about possible recall bias.

> **Recommendation** : Where possible, collect complete histories using retrospective survey questions or administrative data for outcomes that may generate a continuous flow of dollar benefits (e.g., labor market outcomes, public assistance program use, use of costly health or education services).

Third, it may be possible—for some outcomes affected by the SESS intervention—to forecast future benefits beyond the period of follow-up. For example, the cost-benefit studies reviewed in Chapter Four projected future earnings beyond the last follow-up based on the earnings histories of participants observed to date. This allows estimates of increased tax revenue to be projected beyond the last period that participants' outcomes are observed. Likewise, the reduction in future criminal activity and welfare program use was forecast based on observed behavior as of the final follow-up. In other areas, such forecasts may be possible although the methods to do so may require further development. For example, it may be possible to model the link between children's early cognitive gains (e.g., in IQ or achievement tests) and their economic success as adults. We are not aware of any cost-benefit studies that have made such a projection but it should be feasible given other sources of data that would permit estimation of this relationship (see, for example, Currie and Thomas, 1999).

> **Recommendation** : When supported by other empirical evidence, project future benefits based on observed outcomes. Consider additional method development that would permit such forecasts for a broader range of outcomes.

COMPARING COSTS AND BENEFITS OF SESS

The preceding discussion has made it clear that the choice of what type of cost and outcome analysis will be conducted is a driver of the data collection and issues that need to be addressed in preparation of the scorecard. Thus, as indicated by the first recommendation in this chapter, it is important to specify the explicit goals of the cost and outcome analysis in order to determine the nature of the cost and outcome data required. We now briefly summarize the feasibility of undertaking each of the three options outlined at the beginning of this section, given current data collection efforts, and describe some of the changes to data collection that would be required to undertake each of the options in Phase II.

Cost-Benefit or Cost-Savings Analysis

This is the analysis option that would require the greatest modifications to the current data collection plan. This is primarily because under cost-benefit or cost-savings analysis, the analyst would attempt to convert benefits to a monetary value to compare with costs, and the outcomes currently being measured do not lend themselves well to being expressed in monetary value. Hence, to undertake this type of analysis, the types of outcomes collected would need to be expanded as would the duration of the follow-up. Needless to say, cost data would also need to be collected.

This approach would not only take the longest amount of calendar time to execute, as analysis could only get under way after some follow-up time elapsed, but it would also be likely to require the largest budget of the analysis options. This is because new outcomes measures would need to be developed along with a data collection plan for costs. A plan for minimizing participant attrition would need to be devised as well.

This is likely to be the best analysis option only if program sponsors are committed to answering the unique questions addressed by this approach: whether SESS benefits "pay" for their costs, either from the perspective of the government or society as a whole. If this analysis is pursued, it is also important to recognize that the monetary estimates of program benefits are likely to be conservative. Consequently, the program impacts in those domains that can be monetized must be sufficiently large, and sustained over a long enough period, to generate benefits that exceed program costs. The conservative nature of the benefit calculations may produce disappointing results, especially when only short-term results are available. The program may only appear to be cost-beneficial when the evaluation has incorporated information about program outcomes observed a decade or more after the intervention has ended.

Cost-Effectiveness Analysis

Cost-effectiveness analysis for SESS would primarily entail supplementing current data collection with cost data. Not as formidable as the changes required to implement cost-benefit analysis, collecting cost data nevertheless entails large time and resource investments in either or both of the Phase I extension sites and Phase II sites.

This option would answer questions about the relative effectiveness of implementing SESS at PC or EC sites, whether targeting the program to particular participants made a difference, and which treatment components yielded the greatest gains. All of these could be compared on a per-dollar basis if accompanied by cost data.

Replication Analysis

The final type of cost and outcome analysis, which could be undertaken for SESS, is an assessment of the cost of implementing the program in additional sites. This would be most valuable if policymakers envisioned "scaling-up" SESS in the future or if they expected that other agencies might begin to implement the program. If future expansion of the program to other sites is not anticipated, this option has little merit.

This analysis would require collecting cost data, as in the other two analysis options. However, unlike in the cost-benefit or cost-effectiveness options, it would not be particularly important to collect outcome data. It would be important to include program descriptor information, because this would help future sites gauge the comparability of their setting to SESS demonstration sites.

In sum, there is no right or wrong answer to the type of cost and outcome analysis undertaken for SESS. The objectives of the consumers of the analysis dictate the approach taken, which in turn has implications for the collection and analysis of data on program costs and benefits. Clearly, program decisionmakers may have to make tradeoffs between what they might like to achieve and how much of a resource commitment they are willing or able to make.

Chapter Six

CONCLUSIONS

This report has presented an overview of the issues that policymakers would need to assess to be able to select the most appropriate types of cost and outcome analysis for an early childhood intervention program—or to determine whether to even undertake cost and outcome analysis at all. We reviewed the policy scorecard analysis framework used by RAND analysts over the years to organize cost and outcome analysis on a variety of topics. This framework—and the scorecard at its core—helps distinguish between the alternative types of cost and outcome analysis and highlights the data requirements and methodological considerations for the various analysis options.

We also discussed specific methodological issues associated with cost and outcome analysis of early childhood intervention programs and reviewed the results from cost-benefit analysis of three specific programs. Finally, we illustrated the application of cost and outcome analysis methods to the case of the SESS program. Not only does this application address decisions facing that program's stakeholders, it also serves as an illustration of the issues that would need to be considered in assessing the feasibility of undertaking cost and outcome analysis for other early childhood programs.

The recommendations specific to the SESS program evaluation presented in Chapter Five may be restated in more general terms to provide a set of guiding principles regarding cost and outcome analysis of similar types of early childhood intervention programs. These recommendations pertain to evaluation design and the measurement of program costs and benefits. More specifically, we recommend the following:

- Regarding the design of a program evaluation and cost and outcome analysis:
  - Specify the explicit goals of the cost and outcome analysis to guide the scope of cost and benefit data collection and analysis.
  - Identify comparison groups and track the same cost and outcome measures for both comparison and participant groups. If possible, use random assignment to define comparison groups to provide a more valid test of intervention program effects.
  - To minimize attrition in a longitudinal study, devote resources to retaining study subjects.
  - Collect information on program features through site visits and other mechanisms to accurately characterize features of the intervention models as they are implemented and to ensure fidelity to the program model.
- Regarding the collection and analysis of cost data:
  - Collect cost information for both treatment and control groups at each site where the intervention program is implemented.
  - The cost information should be as comprehensive as possible: Costs borne by various parties should be differentiated, the period during which costs are incurred should be identified, and direct and indirect costs, fixed and variable costs, and goods and services provided in-kind should be measured.
  - Plan for proper training and technical support of implementation sites and any cross-site data collection organizations to ensure uniformity in the collection of cost data. Collect information on the cost of data collection, training and support, and the related analyses of the data.

- Regarding the collection and analysis of outcome data:
    - If cost-benefit or cost-savings analysis is the goal, then outcome data should include information for parents and other caregivers in the short term and the long term and for children in the long term in those domains with outcomes that can be readily evaluated in terms of dollars and that can produce large dollar benefits. The choice of specific outcome measures should be guided by findings from related evaluation studies whenever possible.
    - Obtain information from participants that facilitates collection of administrative data and allows effective tracking of individuals to increase response rates at later follow-ups.
    - When possible, collect complete histories using retrospective survey questions or administrative data for outcomes that may generate a continuous flow of dollar benefits (e.g., labor market outcomes, social welfare program use, use of costly health or education services).
    - When supported by other empirical evidence, project future benefits based on observed outcomes. Consider additional method development that would permit such forecasts for a broader range of outcomes.

Although we believe these principles are quite general, ultimately these recommendations should be viewed as guidelines that may need to be tailored to the specific circumstances of a given intervention program and its evaluation design. In the end, the objectives of a program's decisionmakers will dictate the shape of the analysis. As we have seen, cost and outcome analysis is not one method but rather a set of methods, which serve different purposes, place different demands on data collection, and themselves require differing amounts of resources.

The general policy scorecard analysis tools considered in this report, and those specific to cost and outcome analysis, have great promise for improving decisionmaking with respect to investment programs, such as the early childhood interventions represented by SESS and its counterparts. The cost-benefit analyses of the three programs reviewed in Chapter Four have been very influential in providing a justification for devoting resources to interventions with at-risk populations during early childhood. Although results demonstrated for the specific programs, such as the Perry Preschool program, Elmira PEIP, and Chicago CPC, will not necessarily be replicated in other sites implementing the same design or for other program designs, the evidence that program benefits can far outweigh program costs provides proof of the principle that well-targeted investments now can be paid back by future cost savings and benefits to society. When used with skill and judgment, these methods applied to such other programs as SESS will further broaden our base of knowledge with regard to the value of these investments and assist decisionmakers in their choice among program alternatives.

STARTING EARLY STARTING SMART GRANT SITES

This appendix provides additional detail about the *Starting Early Starting Smart* (SESS) grant sites and their programs. The SESS program is an initiative of the Office on Early Childhood, Substance Abuse, and Mental Health Services Administration (SAMHSA) and the Casey Family Programs, along with other federal sponsors. Patricia Salomon, Director of the Office of Early Childhood at SAMHSA, oversees the SESS program along with project officers Michele Basen, Velva Spriggs, and Jocelyn Whitfield, and staff Shakeh Kaftarian. At the Casey Family Programs, the partnership is overseen by Jean McIntosh and Barbara Kelly-Duncan, along with project officers Eileen O'Brien and Peter Pecora.

The SESS program currently operates in 12 sites across the United States Table A.1 lists each of the study sites and the associated principal investigator, project director, and local researcher, first for the primary care (PC) sites and then for the early childhood (EC) sites.[1] Information about the Data Coordinating Center is also provided in Table A.1. A brief description of the program at each site follows the table. Further information about the SESS program is provided in Appendix B and Appendix C and is available from the Casey Family Programs (www.casey.org/projects.htm#sess) and SAMHSA (www.samhsa.gov).

---

[1] One of the original SESS grant sites was unable to continue with the study but made several important contributions to the original design and implementation of the project.

Table A.1

SESS Grant Sites

Study Site	Principal Investigator	Project Director	Local Researcher
Data Coordinating Center			
EMT Associates, Inc., Folsom, Calif., (615) 595-7658	Joel Phillips	J. Fred Springer, Ph.D.	J. Fred Springer, Ph.D.
Primary Care Sites			
Boston Medical Center, Boston, Mass., (617) 414-7433	Carol Seval, R.N., L.M.H.C.	Carol Seval, R.N., L.M.H.C.	Ruth Rose-Jacobs, Sc.D.
The Casey Family Partners, Spokane, Wash., (509) 473-4810	Christopher Blodgett, Ph.D.	Mary Ann Murphy, M.S.	Christopher Blodgett, Ph.D.
University of Miami, Miami, Fla., (305) 243-2030	Connie E. Morrow, Ph.D.	K. Lori Hanson, Ph.D.	Emmalee S. Bandstra, M.D. April L. Vogel, Ph.D.
University of Missouri, Columbia, Mo., (573) 884-2029	Carol J. Evans, Ph.D.	Robyn S. Boustead, M.P.A.	Carol J. Evans, Ph.D.
University of New Mexico, Albuquerque, N.M., (505) 272-3469	Andy Hsi, M.D., M.P.H.	Bebeann Bouchard, M.Ed.	Richard Boyle, Ph.D.
Early Childhood Sites			
Asian American Recovery Services, Inc., San Francisco, Calif., (415) 541-9285, ext. 227	Davis Y. Ja, Ph.D.	Anne Morris, Ph.D.	Anne Morris, Ph.D.
Child Development Inc., Russellville, Ark., (501) 968-6493	JoAnn Williams, M.Ed.	Carol Amundson Lee, M.A., L.P.C.	Mark C. Edwards, Ph.D.
Children's National Medical Center, Washington, D.C., (202) 884-3106	Jill G. Joseph, M.D., Ph.D.	Amy Lewin, Psy.D.	Michelle J. C. New, Ph.D.
Johns Hopkins University, Baltimore, Md., (410) 955-3989	Philip J. Leaf, Ph.D.	Jocelyn Turner-Musa, Ph.D.	Philip J. Leaf, Ph.D.

Table A.1—continued

Study Site	Principal Investigator	Project Director	Local Researcher
Division of Child and Family Services, Las Vegas, Nev., (702) 486-6147	Christa R. Peterson, Ph.D.	Laurel Swetnam, M.A., M.S.	Margaret P. Freese, Ph.D., M.P.H.
The Tulalip Tribes Beda?chelh, Marysville, Wash., (360) 651-3282	Linda L. Jones, B.A.	Linda L. Jones, B.A.	Claudia Long, Ph.D.
The Women's Treatment Center, Chicago, Ill., (773) 373-8670, ext. 302	Jewell Oates, Ph.D.	Dianne Stansberry, B.A., C.S.A.D.P.	Victor J. Bernstein, Ph.D.

PRIMARY CARE GRANT SITES

Boston Medical Center, Department of Pediatrics

Participants: 200.
Population: African-American, Hispanic, and Haitian, ages birth to six months.

Boston Medical Center is a primary care site studying the integration of behavioral health services—Project RISE (Raising Infants in Secure Environments)—into its Pediatric Primary Care Clinic. Project RISE provides integrated services from multiple internal service departments at the medical center and develops referrals to external collaborators. The service integration strategy addresses barriers to access, and families receive transportation to some appointments as necessary. Collaborative agreements have been established with internal departments (e.g., Behavioral Health Services, Center for Excellence in Women's Health, Addiction Service of the Boston Public Health Commission, and River Street Detoxification Center).

The sample population for Project RISE includes inner-city, low-income caregivers who speak English, Spanish, and Haitian Creole and are experiencing a range of risks for mental health and/or substance abuse problems. Participating parents and other caregivers (1) have a history of substance abuse/addiction and/or mental health problems or (2) have active substance abuse/addiction and/or mental health problems or (3) must be considered at-risk stemming from the presence of one or more other risk factors. Parents and other caregivers with major psychotic mental illness are excluded. The control group receives standard pediatric primary care at Boston Medical Center and transportation to regular well-child visits. The randomly assigned intervention and control groups include 100 families each, who are a diverse group of African-American, Haitian, Hispanic, and white non-Hispanic families newly immigrated from 30 different countries. Targeted children are newborn infants. Mother/infant dyads are screened to eliminate serious developmental and health risks (e.g., very low gestational age, HIV positive).

The core intervention team consists of family advocates and behavioral health specialists. Family advocates assigned to each intervention family are central to the Project RISE service strategy. Each family advocate handles case management activities and regularly visits each assigned family at home and in the primary care clinic. Family advocates see families beginning with the first well-child office visit (three to five days old), at age two weeks, and approximately every two months or as needed to age 24 months. They also home visit as needed. They assist the primary care staff in the following up of referrals to specialty clinics within the medical center (e.g., clinics for exposure to lead, failure to thrive). Advocates also work closely with behavioral health specialists (substance abuse, mental health, and child development).

The behavioral health specialists serve as liaisons between pediatrics and internal and external agencies, such as psychiatric inpatient facilities, substance abuse treatment programs, and early intervention programs. They see families as needed, provide assessment and crisis intervention, and facilitate referrals to psychiatric services, substance abuse services, and early intervention by forging collaborative relationships with external agencies. To simplify the referral process for Project RISE parents and caregivers, two behavioral health specialists are assigned to treatment teams in Behavioral Health Services and a third is assigned to Addiction Services.

Casey Family Partners: Spokane

Participants: 170.
Population: 72% white non-Hispanic, 6% African-American, and 22% mixed heritage, ages birth to two and a half years.

Casey Family Partners: Spokane (CFPS) is a primary care site providing assessment and treatment to children and families who have been referred to Child Protective Services (CPS) for child abuse or neglect. Although CFPS serves families affected by both abuse and neglect, only neglect cases are eligible to participate in the SESS study. The target population is 72 percent white non-Hispanic, 6 percent African-American, and 22 percent mixed heritage. The total sample size will be 70 treatment and 100 control children.

The goal of CFPS is to restore children and their families to a healthy, productive life and to expedite permanency planning. A strength-based, intensive case management model is coupled with co-located mental health counseling and substance abuse treatment services, as well as screening and referral for pediatric health, developmental, and parenting skills services.

CFPS case managers ("Family Team Coordinators") work in tandem with CPS social workers assigned to each intervention family to support the family in achieving service goals, while ensuring that the services required for resolving dependency issues are obtained. Family service plans are developed in conjunction with a family team, composed of the client's family, extended family, friends, and collaborators working with the family. The CFPS SESS program focuses on the service needs of both the child and the parent, whereas child welfare decisionmaking typically focuses on the parent's problems that led to the abuse and neglect. Addressing the child's service needs, co-locating critical services in one convenient location, and empowering clients to develop and involve natural support groups of families and friends in their treatment are hallmarks of the CFPS program.

University of Miami School of Medicine's Perinatal CARE Program

Participants: 242.
Population: 52% African-American, 29% Hispanic, 12% Caribbean, and 7% white non-Hispanic, ages birth to three years; 53% of caregivers are known substance users at enrollment.

Miami's Families SESS is administered by the University of Miami (UM) School of Medicine's Perinatal Chemical Addiction Research and Education (CARE) Program. This site is based at the Juanita Mann Health Center (JMHC), a UM/Public Health Trust Community Health Center, which provides a full array of primary health care services in high-risk neighborhoods. The total sample size is 121 intervention children and their families and 121 comparison children and their families.

The Perinatal CARE Program collaborates with various community organizations that provide direct health care, substance abuse treatment/prevention, adult and child mental health, and basic needs services. The JMHC medical staff and Healthy Start High-Risk Children's Program community health nurses are fully integrated into the multidisciplinary team. Collaboration with substance abuse treatment providers has consisted of prioritized referral processes and ongoing consultation with treatment center staff to monitor and support client progress. Simplified referral and co-staffing procedures have been established with several mental health providers. Streamlined referral and service access with early intervention providers has ensured that children identified as developmentally delayed receive immediate evaluation and placement.

Program services include the following:

- *Care Coordination.* Care coordinators, supported by a multidisciplinary team, provide intensive services in a flexible, family-centered format to maintain rapport and facilitate family participation in interventions. Activities include regular face-to-face contact at home visits and on site at JMHC; appointment scheduling, reminders, and follow-up; ongoing needs assessment and participatory family service planning; facilitation of needed service referrals (including basic needs) through cross-agency contacts; and ongoing referral follow-up to assess and address barriers to service utilization.
- *Mental Health and Substance Abuse Treatment and Prevention.* Training for all levels of SESS and collaborating agency staff in the areas of substance abuse and mental health is essential to properly serving families affected by these issues. Ongoing clinical evaluation and informal observation of caregivers' substance use and mental health status is equally important, because these factors are dynamic. SESS staff utilize a flexible approach,

addressing these issues with caregivers at their current level of readiness for change. Crisis intervention and stabilization services are often needed, and treatment engagement efforts are intensive when a need for formal treatment is identified. These engagement activities attempt to overcome treatment barriers through ongoing discussion and supportive encouragement by all SESS staff, solicitation of the support of family members and significant others, and a focus on the impact of parental functioning on children and families. When formal referrals are unwanted or not necessary, short-term individual and family counseling sessions are provided by licensed SESS staff. Preventive educational topic groups related to mental health and substance abuse prevention have been offered monthly on various requested topics.

- *Parenting Interventions.* Several group and individual services are designed to support successful parenting of infants and young children, and efforts are made to include all significant caregivers—mothers, fathers, extended family, and alternative caregivers. Interventions encourage the development and maintenance of appropriate family and peer support systems. Families find it helpful that individual and home-based parenting sessions are available when issues cannot be appropriately addressed in a group setting or they are unable to attend. Two formal group curriculums are described below, and families participate in a formal graduation ceremony following completion of each group. An ongoing grandparents' support group and parent advocacy group meet regularly.

The "Baby & Me" Group is a 14-week parent-infant therapy program that promotes attachment, caregiver knowledge and understanding of infant development and behavior, and empowerment/insight into the impact of the caregiving environment. Each session with three to five parent-infant dyads is two hours and includes group process activities, structured parent-child interaction, practical didactic discussions, and work on a baby book. Didactic topics include attachment, infant communication cues, crying/soothing, sleep/wake patterns, infant medical care, feeding, safety, child abuse prevention, stress management, and anticipatory developmental guidance.

Sessions are designed to facilitate discussion in a manner that is fun and engaging, as well as educational.

The 14-week "Strengthening Multiethnic Families and Communities Program" meets for three hours weekly with 10 to 12 parents. It emphasizes raising children in violence-free environments. Violence prevention is addressed through ethnic/cultural roots, parent-child relationships, parent modeling in the family and community, and parent teaching and discipline. The curriculum helps parents teach children to express emotions, develop empathy, manage anger, and enhance life skills needed to function in society. The program also integrates positive discipline approaches aimed at fostering self-esteem, self-discipline, and social competence. Developing cultural awareness through family rituals/traditions and the importance of community involvement by parents are emphasized.

Curators of the University of Missouri

Participants: 150.
Population: Predominately white non-Hispanic, ages birth to five years.

The University of Missouri is a primary care site studying the integration of behavioral health services into a university pediatric primary care clinic located in Boone County, Missouri. The Healthy Foundations for Families Program serves children between birth and five years of age who live within Boone County. The population served in the pediatric primary care clinic is predominately white non-Hispanic, with a small minority and international population. Referrals are from physicians or self, and selection within the population is based on the caregiver needs with respect to parenting stress. After screening, participants are randomly assigned to the intervention (n = 75) or comparison (n = 75) groups. Those who are not assigned to the intervention receive the usual standard of care, which typically involves referral to other community or hospital-based services from the primary care clinic.

The intervention integrates health and human service professionals working with very young children and families. The professional team includes an on-site recruiter and the child's pediatrician. Family associates are housed in the community. Mutually agreed-on referral forms and release of information forms have been developed to allow for a more expedient and efficient way to initiate the referral/intake process for families. Contracted agencies include those who provide the following:

- Substance abuse counseling.
- Early childhood education.
- Parent education.
- Therapeutic interventions for emotional and behaviorally challenged children and their families.
- Intervention to families with histories of child abuse and neglect.

The family associate is responsible for working with families to identify and coordinate services for the child and family and provide age-appropriate anticipatory guidance from parents in the areas of child health, development, and parent-child interaction. For services beyond those provided at the clinic, families are referred to contracted agencies and other services within the community. To facilitate access to these services, wraparound funds have been established to support program families who experience transportation and child care difficulties. Flexible funds are also available to pay for therapeutic intervention, as well as support services like child safety items, utility bills, or a parenting class.

The community and clinic-based professionals involved receive training on cross-professional issues, culturally competent care, family-centered care of families with young children, anticipatory guidance, and emotional/behavioral problems in young children. In addition, community agencies have been contracted to serve as consultants with regard to barriers that prevent participants from keeping appointments and following through with services.

University of New Mexico

Participants: 200.
Population: Reflects the major ethnic groups in Albuquerque: Hispanic, white non-Hispanic, African-American, Native American, and multiracial, ages birth to three years.

The University of New Mexico Health Sciences Center (HSC) in Albuquerque is the site for the Starting Early to Link Enhanced Comprehensive Treatment Teams (SELECTT) program for families and their children. For the purposes of this study, only families residing in the greater metropolitan area of Albuquerque, within a 40-mile radius, participate in SELECTT.

Families are recruited through referrals from HSC staff, including its specialty clinics and collaborating programs, partner agencies that include private hospitals, Head Start and Early Head Start, and through recruitment presentations made at Career Works/Welfare to Work orientation classes. The program enrolls children under three, with continuing service to age seven, when there is identified family substance use, mental health, domestic violence, and/or unsupported teen issues.

Once a family has been identified as meeting the SELECTT criteria, they are assigned randomly to a treatment group or a control group. Both receive case management services, although those in the control group receive a minimum of four hours of case management per year. Those in the intervention group receive intensive case management, according to a strengths-based, solution-focused approach to engaging and working with families. All service assessment and provision is predicated on the belief that families will become more productive if they focus on healthy behaviors that produce positive change. Families benefit from an interdisciplinary team and case review (i.e., a family service delivery plan), during which service providers discuss goals, identify specific program outcomes, and review family progress in attaining these goals and outcomes.

SELECTT offers child-centered, family-focused services in three locations: at home, in an integrated HSC clinic held one day per week at the Family Practice Clinic of the HSC, or in the SELECTT offices. The unique feature of the program is its capacity to address the needs of the entire family, focusing on healthy behaviors that produce positive change. Program services include the following:

- Primary, Coordinated Medical Care.
- Case Management Services.
- Child Developmental Assessment and Intervention.

- Legal Services.
- Solution Focused Clinical Approaches.
- Substance Use Counseling.
- Mental Health Counseling for Children and Adults.
- Parenting Support Groups.
- Interdisciplinary Team Services.
- Parent Advisory and Community Steering Advisory Committees.
- Extensive Community Referral Base to Early Intervention, Behavioral Health Services.

As a result of its programmatic efforts toward service integration, SELECTT merged with three other programs at the HSC to provide a continuum of services for high-risk children and their families. This collaboration will enhance services across the four programs by offering a wider spectrum of services, cross-training, streamlined documentation, and eventually, a pooling of financial resources.

SELECTT's Steering Committee meets monthly with its HSC and community collaborators to discuss program policy, service issues, and other issues to ensure that services are provided to the families. The principal investigator and program manager are heavily involved in a variety of local and state ad hoc and formal groups, whose goals are to further systems and services integration in specific service areas, such as domestic violence, child witness to violence, early intervention, health care/Medicaid issues, home visiting, and mental health/substance abuse. Among its successes, SELECTT counts its mobilization of the Albuquerque and New Mexico community at its "Community Forum," held in Albuquerque in October 2000, which focused on "Making New Mexico a Child-Friendly State."

EARLY CHILDHOOD GRANT SITES

Asian American Recovery Services, Inc.

Participants: 291.
Population: Predominately Chinese with a minority of Hispanic and African-American, ages three to five years.

Asian American Recovery Services, Inc., is an early childhood grantee assessing the integration of services for an at-risk population composed largely of recently immigrated families. The target population consists of children and their family members at four preschools operated by Wu Yee Children's Services in two inner-city San Francisco neighborhoods. The total sample is 191 intervention children and 100 comparison children. The comparison schools were selected based on their proximity to these neighborhoods, ethnic background, and school size.

Through SESS, the intervention children and their families participate in "CAPS": Comprehensive Asian Preschool Services. The CAPS program is supported by multidisciplinary community partnerships, which include AARS, Inc.; Wu Yee Children's Services; Chinatown Child Development Center (CCDC); and Chinatown Public Health Center. To facilitate organizational collaboration, community partners meet monthly to review policy issues and make progress toward reducing barriers to accessing services.

The CAPS intervention involves both a family advocate and a multidisciplinary case management team. Family advocates provide flexible, responsive, personal contact and support for families. The multidisciplinary family service team, which includes the family advocate, early childhood teaching staff, and a mental health consultant, assesses and plans for service integration for each family. The intervention combines intensive services designed to strengthen family capacity, child development, and access to behavioral health services for assessed families. Children receive enhanced child development services as part of their preschool classes. SESS provides for a partnership with CCDC, a community mental health agency specializing in working with immigrant families. The CCDC mental health consultant provides observation, assessment, and guidance to staff. Children and families in need of additional behavioral health services are referred to community partners off site. Additional intervention strategies include the following:

- Socialization groups for identified children.
- Information and referral for families.
- Parent training and empowerment groups.

- Family relationship enhancement activities.
- Home visiting.

Each year the program operates parent empowerment groups. The program also offers an eight-week, culturally appropriate parent education series at the intervention sites. Parents unable to attend the series receive this information through the family advocates during home visits. The program interventions will continue, according to family need, for up to three years. SESS services are provided at both the early childhood centers and in the home, striving to meet the unique needs of each family.

Child Development, Inc.

Participants: 240.
Population: Primarily white non-Hispanic and African-American, ages three to five years.

Child Development, Inc., is an early childhood site assessing the integration of behavioral health services into Head Start sites serving nine rural Arkansas counties. The intervention and comparison groups consist of children who entered Head Start at age three during the 1998–1999 school year. The sample size is 240—120 intervention children and 120 comparison children. Treatment sites in the target communities were randomly selected, then matched with comparison sites according to center size and type, community income level, number of classrooms, ethnic background of the student body, and age of the Head Start facility. Children at both sites are primarily white non-Hispanic or African-American. Any children who receive parental consent in the intervention and comparison centers are study participants.

The intervention is organized at several levels: community, classroom, and individual family. At the community level, each intervention center has a regional steering committee. The steering committee operates separately from the interdisciplinary team, functioning as a policy organization designed to decrease interorganizational barriers and enhance collaborative capacity. Steering committee members include collaborating agencies, such as the local mental health agency and community mental health providers, the local substance abuse treatment agency, criminal justice, the public school system, county child protective services, victim's assistance, parents, and the Head Start centers. Staff in community organizations receive SESS-sponsored cross-training in such issues as cultural sensitivity in service provision and multiple service coordination. The project conducts extensive training on issues related to resiliency, substance abuse, and child and family issues, focusing on the development of on-site dialogue teams, increased on-site training, and resource enhancement.

At the classroom level, classrooms receive support through training of teachers and staff, and through the provision of behavior management specialists and case managers who assist and advise teachers in addressing behavioral problems in SESS classrooms. They also work closely with mental health practitioners in the development of activities for children.

Families and index children receive an intensive array of services and support during their two years of Head Start and seven months of kindergarten. Case management focuses on developing individualized interventions based on family members needs that have been expressed in the family partnership agreements. Caregivers in the intervention group receive extensive training in parenting through education and support groups, parent-child bonding activities, and the incorporation of prevention activities into parent meetings. Intervention children and families receive most services on site at the Head Start Centers, and home visits provide additional service delivery. Mental health and substance abuse services not co-located on site are made available at collaborating agencies or other referral facilities.

The lead agency provides behavioral health services to intervention children and parent education and training to caregivers. Collaborating agencies provide support groups, mental health services, and outpatient and residential substance abuse services. Collaborating agencies have increased accessibility by extending service hours and simplifying administrative requirements. For families who have difficulty paying for mental health or substance abuse services, the intervention provides a flexible funding source to pay for services, copayments, and deductibles when no other payment sources exist.

Children's National Medical Center

Participants: 280.
Population: 60% Latino, 25% other immigrant, 15% African-American, age four years at recruitment.

The Children's National Medical Center is an early childhood grantee testing the effectiveness of service integration in a Head Start setting in the suburban environment of Montgomery County, Maryland. The sample size is approximately 280—140 intervention and 140 comparison children. Both groups include families and their four-year-old children who attend Head Start. All families whose children attend one of four Head Start schools may participate in the study. Participants are assigned to intervention or comparison groups based on the school attended. Two of the four schools were randomly designated as intervention sites and two as comparison sites. The sample is estimated to be 60 percent Latino, 25 percent other immigrant, and 15 percent African-American.

Intervention provided by SESS staff takes place in the Head Start classrooms and participants' homes. Additional services are delivered in various public and private community agencies. The planned intervention integrates and facilitates access to mental health, substance abuse, educational, physical health, and social services (including housing, financial assistance, vocational training, adult education, and other social service programs).

The collaboration is designed to reduce unmet needs for a variety of mental health, behavioral, and social services through effective service integration of existing community services supplemented by specific home and school-based interventions. Both types of services are provided through linkages to community organizations. The Family Services Agency, Inc. (FSAI), provides regular home visitations by Peer Family Support Workers (FSW) to intervention families to support normative development and effective parenting. FSWs also develop relationships with the family, provide assessments, support family functioning, make recommendations and referrals, assist in follow-through on referrals, and coordinate services. Through Connect for Success (CFS), early childhood mental health specialists provide weekly consultation to Head Start staff in the intervention classrooms. Under the supervision of a clinical psychologist and bilingual MSW, the FSAI and CFS staffs have regular case conference meetings to discuss the needs of specific families, develop intervention plans, and ensure the integration and coordination of home and school interventions.

Service integration and facilitation occurs at multiple project levels. First, representatives from public and private service providers participate in the Montgomery County SESS Community Consortium, which meets regularly to better understand and accomplish service integration. Second, FSWs serve as case managers with intervention families to facilitate access to services and coordinate services used by families with multiple-sector needs. Third, cross-training, particularly in substance abuse and child development, is conducted for SESS, Head Start, and community provider staff. Finally, regular case conferences facilitate multisector integration by addressing the needs of families requiring services from multiple agencies.

The intervention changes significantly in the second year, when the intensity of the home visitation component is reduced and classroom consultation is no longer available. During the second year, the children make the transition into public school kindergarten—a transition that is often a source of stress. The second-year intervention is intended to provide a bridge to independence.

Johns Hopkins University

Participants: 540.
Population: African-American, ages three to five years.

Johns Hopkins University School of Hygiene and Public Health is an early childhood site studying the integration of behavioral health services into two Head Start Centers in Baltimore. The intervention group includes African American children ages three to five and their families, compared with children attending two similar Head Start programs without SESS services. The total sample size is 320 intervention and 220 comparison children. The program is offered to all children and their families at the intervention centers.

The intervention strategy blends preventive services to families with assessment and case management for effectively addressing behavioral health problems potentially impacting the development of index children. All Head Start programs screen children to identify their specific needs and refer them to the appropriate services. However, the intervention group benefits from additional on-site services, including a mental health clinician and resource coordinator who work collaboratively with Head Start staff and community providers to expand and coordinate available services to Head Start children and their families.

Programmatic efforts focus on the following:

- Providing families with services are coordinated on-site and in the community.
- Staff development.
- Parent training.
- Family support groups.

Specifically, an on-site clinician is available to provide direct services to families and staff (staff consultations) and to facilitate family group services. Community-based services are coordinated and integrated through developing a network of services within the community (e.g., substance abuse). At each site, a family community resource coordinator has been added to augment Head Start staff and to work with families and staff to help families access the coordinated services as well as other services they need.

Families have the opportunity to participate in the Pyramid to Success program. This curriculum is designed to help parents develop effective discipline strategies for their children, with a focus on heritage-based and strength-based ways to promote the development of African-American children. In addition, parents have the opportunity to participate in the Families and Schools Together program, a whole-family support group model with an emphasis on substance abuse prevention.

Head Start staff at the two intervention sites participate in joint staff development trainings several times during the school year, as well as site-specific trainings. An advisory group of Head Start parents as well as input from advisory groups from citywide services systems (e.g., Baltimore Substance Abuse Systems) help facilitate the progress of the program.

The on-site clinical services, family parenting/support groups, and staff development activities are delivered in the Head Start Centers. Service integration and coordination activities are coordinated through the Head Start Centers with services received at community-based program sites.

The State of Nevada Division of Child and Family Services

Participants: 192.
Population: Approximately 55% African-American, 35% Hispanic, 10% white non-Hispanic, and a small number of Native American and Asian, ages three to four and a half years.

The state of Nevada is evaluating the impact of New Wish, a project that provides the integration of behavioral health, developmental, substance abuse treatment, and family advocacy services into Head Start sites in Clark County. Targeted children range in age from three to about five years and must be enrolled in Head Start. In Las Vegas, the major city in Clark County, roughly 55 percent of its Head Start preschoolers are African-American, 35 percent are Hispanic, and 10 percent are primarily white non-Hispanic, with a small number of Native Americans and Asians. The study sample size is 192—80 intervention and 112 comparison children. Once families are enrolled in the intervention, services are provided whether or not the child remains in Head Start. The comparison group, which receives traditional Head Start services, is selected from demographically similar Head Start centers. Teachers refer children in need of behavioral health services to the study at both the intervention and comparison centers.

Within the community two powerful barriers to behavioral health and substance abuse treatment programs have been observed: (1) mistrust of formal systems and of individuals who work for them by families who need the programs and (2) fees, transportation, and child care are major issues among the targeted population. New Wish addresses these barriers in the following ways:

- Case managers and family specialists (parent advocates) teach parents to be more effective as advocates and service coordinators.

- Many services are co-located at Head Start centers or provided in families' homes.
- Special arrangements are made to access and support chemical dependency treatment.
- Linkages with collaborators provide access to county mental health services.
- Transportation and childcare are provided as necessary.

The intervention involves the integration of behavioral health services for Head Start children, parents, and families. This includes family and adult mental health (Early Childhood Services, Southern Nevada Adult Mental Health), substance abuse treatment (Bureau of Alcohol and Drug Abuse funded programs in Southern Nevada), developmental services for children (Clark County School District), and family advocacy (Parents Encouraging Parents). Each family chooses a team of representatives from programs providing services to that family. This team meets at least quarterly with parents to formulate a broad-based family intervention plan and to coordinate services. Each family chooses a case manager for the team, who helps parents learn how to achieve follow-through, establish collaboration with service providers, set treatment goals, and achieve them. All service providers communicate changes of plans or difficulties in implementation of service plans with the case manager.

Behavioral health services are offered in the home or at the child's Head Start site by New Wish counselors. More intensive child behavioral health services, such as psychiatric evaluation, medication monitoring, and day treatment, are provided at the most convenient Early Childhood Services site. Developmental services, adult mental health programs, and substance abuse treatment programs are provided by collaborators at the nearest appropriate site. Referrals are expedited for New Wish families.

New Wish counselors are based at New Wish Head Start sites where they are generally available for informal conversation and consultation with parents and teachers. They perform a range of prevention programming for children, adults, and families. Their involvement and usefulness to families results in more openness about families' problem areas.

The Tulalip Tribes

Participants: 201.
Population: Native American, ages three to five years.

The Tulalip Tribes' beda?chelh ("our children") is an early childhood grantee assessing integrated services for "at risk" three- to five-year-old tribal and mainstream children and their families. The Tulalip tribal children and families are accessed through Catholic Community Services' Childspace in Everett and St. Mike's Tikes preschool in Olympia, both of which serve smaller, intact communities within a larger suburban setting. Lummi Head Start provides the comparison for the Tulalip preschools, because Lummi is a Northwest tribal community similar to the Tulalip Tribes. The South Everett Montessori and the South Sound YMCA preschools are comparison sites for the mainstream groups because they serve families socioeconomically similar to those served at ChildSpace and St. Mikes' Tikes. In both tribal and mainstream intervention sites, beda?chelh believes in a mind, body, and spirit approach to reducing risks and enhancing protective factors in children and their parents, and interventions are designed to strengthen individual skills by strengthening the bonds between children and their families and communities. The total sample size is 113 intervention and 88 comparison children.

The intervention involves service integration strategies at the individual, classroom, and community levels. Multidisciplinary teams composed of family members, case managers, child therapists, clinical and legal consultants, child welfare workers, and treatment providers from substance abuse, mental health, and domestic violence fields assess and develop service plans for index children and their families. Interagency collaboration occurs through participation on the multidisciplinary team and on professional advisory boards, which guide the project. Several intercommunity collaborative ties and partnerships extend service provision to the larger communities in which index children reside. All of the above integration strategies are unique to the SESS project, with the exception of the multidisciplinary team. Even this team, however, has been significantly expanded and strengthened under the SESS project.

The integrative mechanisms will guide delivery of and enhance access to services. All index children will receive the following:

- Enhanced preschool curriculums (violence and alcohol, tobacco, and other drug-prevention curriculum through use of the Nee-Kon-Nah Time curriculum).
- Reading readiness and connectedness/bonding through traditional storytelling.
- Milieu therapy in the preschools.
- Gymnastics lessons.

Case management provides access and follow-through for child therapy, mental health services, chemical dependency treatment, family preservation services, domestic violence treatment (for perpetrators and victims), housing assistance, and parenting education and support. These services are provided by the grantee, its partnering agency, and collaborative agencies and organizations.

The curriculum and child-centered services are provided at the early childhood centers and other services are provided at nearby and convenient locations. Family preservation services are provided in the home, as are other services if caregivers are unable to gain access to center-based services. The children and families will receive the majority of their services in the child care/preschool setting. All of the children's enrichment services and the majority of the child therapy are provided in the child care/preschool settings. The family services of substance abuse, mental health, and domestic violence treatment and parenting education occur, for the most part, in the family's small and intact community. Through the project's interagency collaborations, services in the greater community (e.g., inpatient chemical dependency or mental health treatment) are accessed as needed.

The Women's Treatment Center

Participants: 185.
Population: Primarily African-American, ages three to four years.

The Women's Treatment Center is an early childhood grantee collaborating with the Ounce of Prevention Fund and the University of Chicago to study the integration of behavioral health services into a Head Start site located on Chicago's South Side. The intervention group is recruited from two classes and includes African-American children, ages three to four. These children are compared with African-American children receiving traditional Head Start services at a comparison site.

The comparison group is in Head Start, but there are differences in case management procedures. Only the intervention sites receive substance abuse prevention and treatment and mental health services. Both sites have Head Start family support worker services available to them. More intensive family counseling is available at the intervention sites.

The services integration strategy involves the addition of two substance abuse/family support counselors to work directly with all families in the intervention program and additional behavioral health specialists to meet identified needs and make appropriate referrals.

The intervention site receives the following:

- Group parent education.
- Group substance abuse education, screening and referral for treatment and aftercare.
- Mental health screening and referral for treatment.
- On-site family counseling.

A psychologist and a parent-child specialist are available to work with the Head Start staff and family support counselors to develop individual family service plans. These behavioral health specialists are a resource for the integrated staff. On-site substance abuse services for intervention group families are immediately available and free of charge, funded through the SESS grant. Additional service needs are more readily available through the intervention site. Intensive outpatient or residential substance abuse treatment available through the Women's Treatment Center and an outside collaborator provides services for males. Intensive mental health services are provided through an external collaborator. SESS provides for extensive cross-training of professionals from other disciplines regarding the identification, signs, and symptoms of substance abuse.

The bulk of services takes place at the Head Start centers, while such specialized needs as substance abuse treatment take place at the Women's Treatment Center and other collaborating agencies. Each intervention and comparison center has the benefit of a Head Start Parent Advisory Council.

Appendix B

SESS PROGRAM ACKNOWLEDGMENTS

The families and grantees of *Starting Early Starting Smart* (SESS) would like to acknowledge:

Nelba Chavez, Ph.D.
Administrator, SAMHSA
Rockville, Maryland

Ruth Massinga, M.S.
President and CEO, Casey
Family Programs
Seattle, Washington

along with the Casey Board of Trustees and the three SAMHSA Centers—Center for Substance Abuse Prevention, Center for Substance Abuse Treatment, and Center for Mental Health Services—for their vision and commitment to reaching families with very young children affected by environments of substance abuse and mental disorders. Without their innovative public-private partnership and unprecedented support, this initiative would have been impossible.

We further acknowledge the early guidance and program development from Stephania O'Neill, M.S.W., Rose Kittrell, M.S.W., Hildy (Hjermstad) Ayers, M.S.W., Karol Kumpfer, Ph.D., Sue Martone, M.P.A., and Jeanne DiLoreto, M.S.

Many thanks to the SAMHSA-Casey team for their tenacious efforts and unprecedented collaboration:

Joe Autry, M.D.
Acting Administrator
SAMHSA

Pat Salomon, M.D.
Michele Basen, M.P.A.
Velva Springs, M.S.W.
Jocelyn Whitfield, M.A.

Jean McIntosh, M.S.W.
Executive Vice President
Casey Strategic Planning and Development

Barbara Kelley Duncan, M.Ed.
Peter Pecora, Ph.D.
Eileen O'Brien, Ph.D.

MISSION STATEMENTS OF THE NATIONAL COLLABORATORS

SUBSTANCE ABUSE AND MENTAL HEALTH SERVICES ADMINISTRATION (SAMHSA)

SAMHSA's mission within the nation's health system is to improve the quality and availability of prevention, treatment, and rehabilitation services to reduce illness, death, disability, and cost to society resulting from substance abuse and mental illness.

SAMHSA's mission is accomplished in partnership with all concerned with substance abuse and mental illness. SAMHSA exercises leadership in

- eliminating the stigma that impedes prevention, treatment, and rehabilitation services for individuals with substance abuse;
- developing, synthesizing, and disseminating knowledge and information to improve prevention, treatment, rehabilitation services, and improving the organization, financing, and delivery of these services;
- providing strategic funding to increase the effectiveness and availability of services;
- promoting effective prevention, treatment, and rehabilitation policies and services;
- developing and promoting quality standards for service delivery;
- developing and promoting models and strategies for training and education;

- developing and promoting useful and efficient data collection and evaluation systems; and
- promoting public and private policies to finance prevention, treatment, and rehabilitation services so that they are available and accessible.

For more information, visit SAMHSA's Web site at www.SAMHSA.gov.

CASEY FAMILY PROGRAMS

The mission of Casey Family Programs is to support families, youth, and children in reaching their full potential. Casey provides an array of permanency planning, prevention, and transition services, such as long-term family foster care, adoption, kinship care, job training, and scholarships.

The program aims to improve public and private services for children, youth, and families impacted by the child welfare system, through advocacy efforts, national and local community partnerships, and by serving as a center for information and learning about children in need of permanent family connections.

Casey Family Programs is a Seattle-based private operating foundation, established by Jim Casey, founder of United Parcel Service (UPS), in 1966. The program has 29 offices in 14 states and Washington, D.C. For more information, visit our Web site at www.casey.org.

BIBLIOGRAPHY

Adler, M. D., & Posner, E. A. (2000). "Introduction," to cost-benefit analysis: Legal, economic, and philosophical perspectives. A Conference Sponsored by the John M. Olin Foundation and The University of Chicago Law School, *Journal of Legal Studies, 29* (2, part 2), 837–842.

Arrow, K. J. (1951). *Social choice and individual values.* New York: Wiley.

Barnett, S. W. (1993). Benefit-cost analysis of preschool education: Findings from a 25-year follow-up. *American Journal of Orthopsychiatry, 63* (4), 500–508.

_____ (1995). Long-term effects of early childhood programs on cognitive and school outcomes. *The Future of the Children, 5,* 25–50.

_____ (1996). *Lives in the balance: Age-27 benefit-cost analysis of the High/Scope Perry Preschool Program.* Ypsilanti, MI: High/Scope Press.

Benjamin, R., Caroll, S. J., Jacobi, M., Krop, C. S., & Shires, M. A. (1993). *The redesign of governance in higher education.* Santa Monica, CA: RAND, MR-222-LE.

Campbell, F. A., & Ramey, C. T. (1994). Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low-income families. *Child Development, 62* (2), 684–689.

Cannon, J. S., Karoly, L. A., & Kilburn, M. R. (2001). *Directions for cost and outcome analysis of* Starting Early Starting Smart*: Summary of a cost expert meeting.* Santa Monica, CA: RAND, CF-161-TCFP.

Caulkins, J. P., Rydell, C. P., Schwabe, W., & Chiesa, J. R. (1997). *Mandatory minimum drug sentences: Throwing away the key or the taxpayers' money?* Santa Monica, CA: RAND, MR-827-DPRC.

Caulkins, J. P., Rydell, C. P., Everingham, S. S., Chiesa, J. R., & Bushway, S. (1999). *An ounce of prevention, a pound of uncertainty: The cost-effectiveness of school-based drug prevention programs.* Santa Monica, CA: RAND, MR-923-RWJ.

Caulkins, J. P. (2000). Measurement and analysis of drug problems and drug control efforts. In D. Duffee, D. McDowall, L. G. Mazerolle, & S. D. Mastrofski (Eds.), *Criminal justice 2000: Volume 4, measurement and analysis of crime and justice* (pp. 391–449). Washington, DC: USGPO.

Center for Substance Abuse Treatment (CSAT), NEDS (August 1999), The costs and benefits of substance abuse treatment: Findings from the national treatment improvement evaluation study (NTIES), College Park, MD.

Cohen, M. A. (2000). Measuring the costs and benefits of crime and justice. In D. Duffee, D. McDowall, L. G. Mazerolle, & S. D. Mastrofski (Eds.), *Criminal justice 2000: Volume 4, measurement and analysis of crime and justice* (pp. 263–316). Washington, DC: USGPO.

Cook, P. J. (1983). Costs of crime. In S. H. Kadish (Ed.), *Encyclopedia of crime and justice.* New York: Free Press.

Currie, J. (forthcoming). Early childhood intervention programs: What do we know?, *Journal of Economic Perspectives.*

Currie, J., & Thomas, D. (1995). Does Head Start make a difference? *American Economic Review, 85* (3), 341–364.

_____ (February 1999). Early test scores, socio-economic status, and future outcomes. National Bureau of Economic Research Working Paper No. 6943.

Dewar J. A., Builder, C. H., Hix, W. M., & Levin, M. H. (1993). *Assumption-based planning: A planning tool for very uncertain times.* Santa Monica CA: RAND, MR-114-A.

Foster, E. M., & Bickman, L. (2000). Refining the costs analyses of the Fort Bragg evaluation: The impact of cost offset and cost shifting. *Mental Health Services Research, 2* (1), 13–25.

Frank, R. H. (2000). Why is cost-benefit analysis so controversial? *The Journal of Legal Studies, 29* (2, part 2), 913–930.

Goeller, B. F., et al. (1973). San Diego Clean Air project: Summary report. Santa Monica, CA: RAND, R-1362-SD.

Goeller, B. F., et al. (1977). Protecting an estuary from floods—a policy analysis of the Oosterschelde. Leiden, the Netherlands: RAND Europe, R-2121/1-NETH.

Goeller, B. F., & the Pawn Team (1985). Planning the Netherlands' water resources. *Interfaces, 15* (1), 1–33.

Gold, M. R., Siegel, J. E., Russell, L. B., & Weinstein M. (1996). *Cost-effectiveness in health and medicine.* New York: Oxford University Press.

Gramlich, E. M. (1981), *Benefit-cost analysis of government programs.* Englewood Cliffs, NJ: Prentice Hall.

Greenberg, D. F. (1990). The cost-benefit analysis of imprisonment. *Social Justice, 17* (4), 49–75.

Greenwood, P. W., Model, K. E., Rydell, C. P., & Chiesa, J. R. (1998). *Diverting children from a life of crime: Measuring costs and benefits.* Santa Monica CA: RAND, MR-699-1-UCB/RC/IF.

Guralnick, M. J. (Ed.) (1997). *Effectiveness of early intervention.* Baltimore: Paul Brookes Publishing.

Hahn, R. W. (2000). State and federal regulatory reform: A comparative analysis. *Journal of Legal Studies, 29* (2, part 2), 873–912.

Hargreaves, W. A., Shumway, M., Hu, T.-W., & Cuffel, B. (1998). *Cost-outcome methods for mental health.* San Diego, CA: Academic Press.

Harwood, H., Fountain, D., & Livermore G. (1998). *The economic costs of alcohol and drug abuse in the United States, 1992.* Washington, DC: U.S. Department of Health and Human Services.

Hillestad, R. J., Walker, W. E., Carillo, M. J., Bolton, J. G., Twaalfhoven, P., & van de Riet, O. (1996). *FORWARD—freight options for road, water, and rail for the Dutch: Final report.* Leiden, the Netherlands: RAND Europe, MR-736-EAC/VW.

Kamlet, M. (1992). *The comparative benefits modeling project: A framework for cost-utility analysis of governmental health care programs,* Washington, DC: U.S. Department of Health and Human Services.

Karoly, L. A. (forthcoming). Investing in the future: Reducing poverty through human capital investments. In S. Danziger, R. Haveman, & B. Wolfe (Eds.), *Understanding poverty in America: Progress and problems.* Cambridge, Mass.: Harvard University Press.

Karoly, L. A., Greenwood, P. W., Everingham, S. S., Houbé, J., Kilburn, M. R., Rydell, C. P., Sanders, M., & Chiesa, J. R. (1998). *Investing in our children: What we know and don't know about the costs and benefits of early childhood interventions.* Santa Monica, CA: RAND, MR-898.

Keeler, E. B., & Cretin, S. (1983). Discounting of life-saving and other nonmonetary effects. *Management Science, 29,* 300–306.

Keeney, R. L. (1992). *Value-focused thinking: A path to creative decisionmaking.* Cambridge, MA: Harvard University Press.

Keeney, R. L., & Raiffa, H. (1976). *Decisions with multiple objectives: Preferences and value trade-offs.* New York, NY: Wiley.

Kitzman, H., Olds, D. L., Henderson, C. R., et al. (1997). Effect of prenatal and infancy home visitation by nurses on pregnancy outcomes, childhood injuries, and repeated childbearing: A randomized controlled trial. *Journal of the American Medical Association, 278* (8), 644–652.

Klaus, P. A. (1994). *The cost of crime to victims: Crime data brief*, Washington, DC: U.S. Department of Justice, Bureau of Justice Statistics, NCJ 145865.

Larkey, P., Kadane, J. B., Austin, R., & Zamir, S. (1997). Skill in games. *Management Science, 43* (5), pp. 596–609.

Lazar, I., & Darlington, R. (1982). *Lasting effects of early education: A report from the consortium for longitudinal studies*, Monographs of the Society for Research in Child Development, 47 (2-4), Serial No. 195.

McLellan, A. T., Woody, G. E., Metzger, D., et al. (1996). Evaluating the effectiveness of addiction treatments: Reasonable expectations, appropriate comparisons. *Milbank Quarterly, 74* (1), 51–85.

Miller, T. R., Cohen, M. A., & Wiersema, B. (1996). *Victim costs and consequences: A new look*. Washington, DC: U.S. Department of Justice, National Institute of Justice, NCJ 155282.

Mishan, E. J. (1998). *Cost-benefit analysis: An informal introduction*, 4th Edition, London: Unwin Hyman.

NICHD Early Child Care Research Network (1997). Poverty and patterns of child care. In G. J. Duncan & J. Brooks-Gunn (Eds.), *Consequences of growing up poor* (pp. 100–131). New York: Russell Sage Foundation.

Olds, D. L., Eckenrode, J., Jr., Henderson, C. R., et al. (1997). Long-term effects of home visitation on maternal life course, child abuse and neglect, and children's arrests: Fifteen year follow-up of a randomized trial. *Journal of the American Medical Association, 278* (8), 637–643.

Olds, D. L., Henderson, C. R., Jr., Tatelbaum, R., Chamberlin, R., et al. (1986a). Improving the delivery of prenatal care and outcomes of pregnancy: A randomized trial of nurse home visitation. *Pediatrics, 77* (1), 16–28.

_____ (1986b). Preventing child abuse and neglect: A randomized trial of nurse home visitation. *Pediatrics, 78* (1), 65–78.

Olds, D. L., Henderson, C. R., Jr., & Kitzman, H. (1994). Does prenatal and infancy nurse home visitation have enduring effects on qualities of parental caregiving and child health at 25 to 50 months of life? *Pediatrics, 93* (1), 89–98.

Olds, D. L., Henderson, C. R., Jr., Phelps, C., Kitzman, H., & Hanks, C. (February 1993). Effect of prenatal and infancy nurse home visitation on government spending. *Medical Care, 31* (2), 155–174.

Park, G., & Lempert, R. (1998). *The class of 2014: Preserving access to California higher education*. Santa Monica, CA: RAND, MR-971-EDU.

Posner, R. A. (2000). Cost-benefit analysis: Definition, justification, and comment on conference papers. *The Journal of Legal Studies, 29* (2, part 2), 1153–1177.

Quade, E. S. (1989). *Analysis for public decisions* (3rd ed., revised by Grace M. Carter). Englewood Cliffs, NJ: Prentice Hall, 1989.

Raiffa, H. (1968). *Decision analysis.* Reading, MA: Addison-Wesley.

Reynolds, A. J. (1994). Effects of a preschool plus follow-on intervention for children at risk. *Developmental Psychology, 30* (6), 787–804.

_____ (2000). *Success in early intervention: The Chicago Child-Parent Centers.* Lincoln, NE: University of Nebraska Press.

Reynolds, A. J., Chang, H., & Temple, J. A. (April 1997). Early educational intervention and juvenile delinquency: Findings from the Chicago Longitudinal Study. Paper presented at the SRCD Seminar on Early Intervention Effects on Delinquency and Crime, Washington, DC.

Reynolds, A. J., Mann, E., Miedel, W., & Smokowski, P. (1997). The state of early childhood intervention: Effectiveness, myths, and realities, new directions, *Focus, 19* (1), 5–11.

Reynolds, A. J., & Temple, J. A. (1995). Quasi-experimental estimates of the effects of a preschool intervention. *Evaluation Review, 19* (4), 347–373.

Reynolds, A. J., Temple, J. A., Robertson, D. L., & Mann, E. A. (March 30, 2000). Long-term benefits of participation in the Title I Chicago Child-Parent Centers. Paper presented at the biennial meeting of the Society for Research on Adolescence, Chicago.

Rice, D. P., Kelman, S., Miller, L. S., et al. (1990). *The economic costs of alcohol and drug abuse and mental illness: 1985.* San Francisco: University of California, Institute for Health and Aging.

Richardson, H. S. (2000). The stupidity of the cost-benefit standard. *The Journal of Legal Studies, 29* (2, part 2), 971–1003.

Rydell, C. P., & Everingham, S. S. (1994). *Controlling cocaine: Supply versus demand programs.* Santa Monica, CA: RAND, MR-331-ONDCP/A/DPRC.

Schweinhart, L. J., Barnes, H. V., & Weikart, D. P. (1993). *Significant benefits: The High/Scope Perry Preschool Study through age 27.* Ypsilanti, MI: High/Scope Educational Research Foundation, Monographs of the High/Scope Educational Research Foundation, Number 10.

Schweinhart, L. J., & Weikart, D. P. (1980). *Young children grow up: The effects of the Perry Preschool Program on youths through age 15.* Ypsilanti, Mich.: High/Scope Educational Research Foundation, Monographs of the High/Scope Educational Research Foundation, Number 7.

Sen, A. (2000). The discipline of cost-benefit analysis. *The Journal of Legal Studies, 29* (2, part 2), 931–952.

Stein, L. I., & Test, M. A. (1980). Alternative to mental hospital treatment: I. Conceptual Model, treatment program, and clinical evaluation. *Archives of General Psychiatry, 37,* 392–397.

Sturm, R., Gresenz, C. R., Pacula, R. L., & Wells, K. B. (1999). Labor force participation by persons with mental illness. *Psychiatric Services, 50* (11), 1407.

Test, M. A., & Stein, L. I. (1980). Alternative to mental hospital treatment: III. Social cost. *Archives of General Psychiatry, 37,* 409–412.

Tragler, G., Caulkins, J. P., & Feichtinger, G. (forthcoming). Optimal dynamic allocation of treatment and enforcement in illicit drug control. *Operations Research*, 49, (3).

Trumbell, W. N. (1990). Who has standing in cost-benefit analysis? *Journal of Policy Analysis and Management*, 9, 201–218.

Walker, W., et al. (1993). *Investigating basic principles of river dike improvement: Safety analysis, cost estimation, and impact assessment.* Leiden, the Netherlands: RAND Europe, MR-143-EAC/VW.

Walker, W., Poyhonen, M., de Jong, J. H., & van der Tak, C. (1999). *POLSSS—policy for sea shipping safety: Executive summary.* Leiden, the Netherlands: RAND Europe, MR-1043-RE/VW.

White, K. R. (1985). Efficacy of early intervention. *The Journal of Special Education, 19* (4), 401–416.

Yates, B. T. (1996). *Analyzing costs, procedures, processes, and outcomes in human services,* Applied Social Research Methods Series, Volume 42, Thousand Oaks, CA: Sage Publications.

Yoshikawa, H. (Winter 1995). Long-term effects of early childhood programs on social outcomes and delinquency. *The Future of Children, 5*, 51–75.

The SESS Sites

Miami's Families: Starting Early Starting Smart	Florida
Raising Infants in Secure Environments	Massachusetts
Healthy Foundations for Families	Missouri
Starting Early to Link Enhanced Comprehensive Treatment Teams	New Mexico
Casey Family Partners	Washington
National Association for Families and Addiction Research and Education	Illinois*
Child Development, Inc.	Arkansas
Asian American Recovery Services, Inc.	California
Locally Integrated Services in Head Start	Washington, D.C.
Starting Early Starting Smart Head Start Collaboration Project	Illinois
Baltimore BETTER Family and Community Partnership	Maryland
New Wish	Nevada
Beda?chelh Tulalip Tribes Early Intervention in Tribal and Mainstream Communities	Washington
Evaluation, Management and Training, Inc.**	California

*One of the original SESS sites was unable to continue with the study, but it was an important contributor to the original design and implementation of this project. Our thanks to Dr. Linda Randolph and Dr. Ira Chasnoff.

**Data Coordinating Center

Please feel free to be a "copy cat" by making all the copies you want of the entire document; or if sections are copied, please provide the full citation to the report.

For further information on alcohol, tobacco, and illicit drugs, call 1-800-729-6686, 301-468-2600, or TDD 1-800-487-4889. Or visit the World Wide Web at *www.health.org*.

BEST COPY AVAILABLE

SAMHSA

CASEY
FAMILY
PROGRAMS
FOUNDATIONS FOR THE FUTURE

STARTING
EARLY
STARTING
SMART

*Healthy Minds, Healthy
Behaviors, Promising
Lives Right From the Start*

BEST COPY AVAILABLE

CMHS
Center for Mental
Health Services
SAMHSA

CSAT
Center for Substance
Abuse Treatment
SAMHSA

CSAP Center for
Substance Abuse
Prevention
Substance Abuse and Mental
Health Services Administration

RAND

EXECUTIVE SUMMARY

P_

SESS

STARTING EARLY STARTING SMART

Early

Assessing Costs and Benefits of Early Childhood Intervention Programs

Overview and Application to the Starting Early Starting Smart Program

Lynn A. Karoly
M. Rebecca Kilburn
James H. Bigelow
Jonathan P. Caulkins
Jill S. Cannon
James R. Chiesa

BEST COPY AVAILABLE

Smart

ABOUT STARTING EARLY STARTING SMART

Starting Early Starting Smart (SESS) is a knowledge development initiative designed to

- Create and test a new model for providing integrated behavioral health services (mental health and substance abuse prevention and treatment) for young children (birth to 7 years) and their families; and to
- Inform practitioners and policymakers of successful interventions and promising practices from the multi-year study, which lay a critical foundation for the positive growth and development of very young children.

In October 1997, with initial funding of $30 million, the Substance Abuse and Mental Health Services Administration (SAMHSA) and Casey Family Programs embarked on a precedent-setting public/private collaboration. Twelve culturally diverse grantee organizations were selected. Each provides integrated behavioral health services in community-based early childhood settings—such as Child Care, Head Start and Primary Care Clinics—where young families customarily receive services for children. Critical to this project is the required collaboration among funders, grantees, consumers, and local site service providers. Implicit in the design of this project is sustainability planning for secured longevity of the programs.

The SESS approach informs policy-making for:

- Service system redesign
- Strengthening the home environment
- Using culture as a resource in planning services with families
- Service access and utilization strategies
- Targeting benefits for children
- Working with families from a strengths-based perspective

The Research Design

The 12 grantees, working collaboratively, designed a study whereby integrated behavioral health services are delivered in typical early childhood settings. Each site has an intervention and comparison group, and each site delivers similar targeted, culturally-relevant, interventions for young children and their families. A collaboratively determined set of outcomes has been established to evaluate project effectiveness:

- Access to and use of services
- Social, emotional, and cognitive outcomes for children
- Caregiver-child interaction outcomes
- Family functioning

The goal of the SESS research is to provide rigorous scientific evidence concerning whether children and families participating in SESS programs achieve better access to needed services and better social, emotional, cognitive, and behavioral health outcomes than do the children and families not receiving these services. SESS programs may also generate information about opportunities, practices, and barriers to sought-after outcomes. This information is critical to achieving effective public policies.

SESS Extended

It was clear from the early days of SESS that whatever effects were uncovered, longitudinal extension of the study would be valuable. In 2001, SAMHSA and Casey Family Programs embarked upon an extension phase, which will increase understanding of the impact of early intervention as young children enter preschool and school years, when babies or toddlers are asked to meet escalating emotional and cognitive demands. This longitudinal extension can validate early methods and findings and assess their durability. It is anticipated that this work will include additional data points of a refined instrument set and intervention package with the addition of study questions related to cost and value, and other special studies. Additional future plans include applying and validating early SESS lessons learned, key concepts, components, and principles to new settings that serve families with young children.

Summation

In sum, SESS reflects the growing acknowledgement that it is important to target positive interventions to very young children. The infant and preschool years lay a critical foundation for later growth and development. Second, successful interventions for very young children must meet the multiple behavioral health, physical health, and educational needs of families. Third, integrated behavioral health services must be made more accessible to families with multiple needs, which are difficult to meet in a fragmented service system.

For more information about *Starting Early Starting Smart* and related SAMHSA-Casey products, contact *www.casey.org* or *www.samhsa.gov* (SESS section under construction).

BEST COPY AVAILABLE

Assessing Costs and Benefits of Early Childhood Intervention Programs

Overview and Application to the Starting Early Starting Smart Program

EXECUTIVE SUMMARY

Lynn A. Karoly
M. Rebecca Kilburn
James H. Bigelow
Jonathan P. Caulkins
Jill S. Cannon
James R. Chiesa

BEST COPY AVAILABLE

Spring 2001

This report would not have been possible without the contributions of staff from the Office on Early Childhood, SAMHSA, U.S. Department of Health and Human Services, the Casey Family Programs, the Starting Early Starting Smart principal investigators, project directors and researchers, and the parent representatives, who helped design and supervise the data collection. The content of this publication does not necessarily reflect the views or policies of the U.S. Department of Health and Human Services, or the Casey Family Programs, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. government. Responsibility for the content of this report, however, rests solely with the named authors.

Library of Congress Cataloging-in-Publication Data

Assessing costs and benefits of early childhood intervention programs : overview and applications to the Starting Early Starting Smart program / Lynn A. Karoly ... [et al.].
p. cm.
"MR-1336-CFP"—T.p., v. 1.
"MR-1336/1-CFP"—T.p., v. 2.
Includes bibliographical references.
Contents: [v. 1. without special title]—[v. 2] Executive summary.
ISBN 0-9708278-1-4 (v. 1)—ISBN 0-9708278-0-6 (v. 2)—Casey
ISBN 0-8330-2973-8 (v. 1)—ISBN 0-8330-2974-6 (v. 2)—RAND
1. Starting Early Starting Smart (Program) 2. Child welfare—United States—Evaluation. 3. Early childhood education—United States—Evaluation. 4. Child mental health services—United States—Evaluation. 5. Children—Drug use—United States—Evaluation. I. Karoly, Lynn A., 1961– II. Starting Early Starting Smart (Program).

HV741 .A86 2001
362.7'0973—dc21

2001019040

Any or all portions of this document may be reproduced with proper citation: "Source: Karoly, L., Kilburn, R., Bigelow, J. H., Caulkins, J. P., and Cannon, J. S. (2001). *Assessing Costs and Benefits of Early Childhood Intervention Programs: Overview and Applications to the* Starting Early Starting Smart *Program.* Publishers: Seattle: Casey Family Programs; Santa Monica: RAND."

Published 2001 by Casey Family Programs
1300 Dexter Avenue North, Suite 300
Seattle, WA 98109
Telephone: (206) 282-7300; Fax (206) 378-4619;
Internet: www.Casey.org

RAND
1700 Main Street, P.O. Box 2138, Santa Monica, CA 90407-2138
1200 South Hayes Street, Arlington, VA 22202-5050
201 North Craig Street, Suite 102, Pittsburgh, PA 15213
Internet: www.rand.org/
To order from RAND, contact Distribution Services:
Telephone: (310) 451-7002; Fax: (310) 451-6915;
E-mail: order@rand.org

RAND is a nonprofit institution that helps improve
policy and decisionmaking through research and analysis.
RAND® is a registered trademark.

PREFACE

The increased interest in the potential for early childhood intervention programs to save dollars in the long run has focused attention on the potential for cost-benefit and related analyses to aid decisionmakers in their policy choices. The goal of this report is to identify the conceptual and methodological issues associated with the analysis of costs and outcomes of early intervention programs in general and to make recommendations regarding the application of these tools for subsequent demonstration studies of a particular intervention program: *Starting Early Starting Smart* (SESS).

SESS is a public-private collaboration designed to test the effectiveness of integrating behavioral health services within primary care and early childhood service settings for children from birth to age seven. The SESS program is an initiative of the Office on Early Childhood, Substance Abuse and Mental Health Services Administration (SAMHSA), and the Casey Family Programs, along with several other federal sponsors. The program currently operates in 12 sites across the United States and is entering the third year of its first five-year phase. An outcomes evaluation is built into the first phase.

Program sponsors are beginning to plan for a second phase, the design of which they hope will be informed by the first phase. It was during the initiation of this planning process that program sponsors identified a need for cost information to supplement their outcomes information. Recognizing that the literature offered somewhat limited guidance on the specifics of cost considerations in this context, they requested that RAND not only present them with a summary of research bearing on their problem but that we also examine their program and make specific recommendations regarding how cost and outcome analysis could improve their decisionmaking.

This project began with a meeting of cost and outcome analysis experts held in August 2000, convened by RAND on behalf of the Casey Family Programs and the Office on Early Childhood, SAMHSA. Participants at the meeting included four national experts in cost and outcome analysis with backgrounds in mental health and substance abuse, as well as several RAND staff members with experience in cost and outcome analysis. Also participating were staff from SAMHSA, the Casey Family Programs, the SESS Data Coordinating Center, and two of the SESS program sites. The proceedings from the meeting are summarized in the following document:

Cannon, Jill S., Lynn A. Karoly, and M. Rebecca Kilburn, *Directions for Cost and Outcome Analysis of* Starting Early Starting Smart: *Summary of a Cost Expert Meeting*, CF-161-TCFP, Santa Monica, California: RAND, 2001.

Readers interested in more detail are urged to obtain a copy of the full report, *Assessing Costs and Benefits of Early Childhood Intervention Programs: Overview and Application to the* Starting Early Starting Smart *Program*, by Lynn A. Karoly, M. Rebecca Kilburn, James H. Bigelow, Jonathan P. Caulkins, and Jill S. Cannon, which can be obtained from RAND (www.rand.org), the Casey Family Programs (www.casey.org/projects.htm#sess), or SAMHSA (www.samhsa.gov).

This research is funded by the Casey Family Programs. The opinions expressed and conclusions drawn in this report are the responsibility of the authors and do not represent the official views of the Casey Family Programs, SAMHSA, other agencies, or RAND.

CONTENTS

Preface iii

Figures vii

Tables ix

Acknowledgments xi

Acronyms xiii

Executive Summary 1
  The Cost and Outcome Analysis Framework 2
  Applying the Framework to Early Childhood Interventions 5
  Some Illustrative Analyses 7
  Framing a Policy Scorecard Analysis for a Specific Program 9
  Conclusions 12

Appendix

A. *STARTING EARLY STARTING SMART* GRANT SITES 15
B. SESS PROGRAM ACKNOWLEDGMENTS 17
C. MISSION STATEMENTS OF THE NATIONAL COLLABORATORS 19

References 21

FIGURES

S.1 Some Early Childhood Interventions Have Been Shown to Have High Benefit-Cost Ratios 8

TABLES

S.1 Early Childhood Intervention Program Benefit Domains and Illustrative Measures 6
A.1 SESS Grant Sites 15

ACKNOWLEDGMENTS

We thank our sponsors at the Casey Family Programs and the Substance Abuse and Mental Health Services Administration (SAMHSA) for their support of this project. We especially acknowledge the valuable guidance we received from Peter Pecora at the Casey Family Programs, and Michelle Basen and Patricia Salomon from the Office on Early Childhood at SAMHSA. We also benefited from discussions with a number of cost and outcome analysis experts who specialize in mental health and substance abuse interventions. In particular, we thank Anthony Broskowski (President, Pareto Solutions), William A. Hargreaves (University of California, San Francisco), Brenda Reiss-Brennan (President, Primary Care Family Therapy Clinics, Inc.), and Brian T. Yates (American University). These four individuals served as expert panelists for a two-day meeting convened by RAND in August 2000 to discuss the potential use of cost-benefit analysis or related methods in the analysis of subsequent knowledge development and application studies of the *Starting Early Starting Smart* (SESS) program, a public-private initiative funded and directed through SAMHSA and the Casey Family Programs. This document was shaped by the insights they offered during and after the meeting. In addition, we benefited greatly from information about the current SESS evaluation provided by Fred Springer and his colleagues at EMT Associates, which serves as the SESS Data Coordinating Center. The perspectives on the SESS program offered by David Deere and Karen Rossmaier from the Russellville, Arkansas, site, and Miriam Escobar and K. Lori Hanson from the Miami, Florida, site were also insightful. Other participants at the meeting from SAMHSA and the Casey Family Programs also provided useful input.

We are also grateful for the careful technical reviews provided by our RAND colleagues Carole Roan Gresenz and Shin-Yi Wu. Other valuable comments on the draft report were provided by Richard Boyle and Mark Friedman. Finally, we also thank Patrice Lester and Claudia Szabo of RAND for their very able assistance with the assembly and production of the document.

ACRONYMS

EC	Early childhood (program)
PC	Primary care (program)
PEIP	Prenatal/Early Infancy Project
SAMHSA	Substance Abuse and Mental Health Services Administration
SESS	*Starting Early Starting Smart* (program)

EXECUTIVE SUMMARY

Agency and program administrators and decisionmakers responsible for implementing early childhood intervention programs are becoming more interested in quantifying the costs and benefits of such programs. Part of the reason for this is that foundations and other funders are putting more emphasis on results-based accountability. At the same time, arguments for the value of early childhood intervention are being made within the public sphere on the basis of published estimates of costs and benefits. Program implementers are naturally attracted by statements that a certain intervention produces $4 in savings for every $1 it costs and would like to make similar statements about their own programs. Meanwhile, decisionmakers without particular interest in any given program would like more quantitative decision aids when it comes time to choose among a variety of possible program models or program improvements to implement.

Our objective here is to offer assistance to decisionmakers and program implementers considering an assessment of costs and outcomes. We do not offer a specific step-by-step manual, but we discuss the kinds of issues that must be taken into account and why. We do so in enough detail that readers can decide if this type of quantitative analysis is the right course for them and, if so, can knowledgeably interact with an expert cost-outcome analyst. While we understand that some readers will want to undertake analysis of costs and outcomes to justify a program in which they have a special interest, we take the viewpoint here of an unbiased allocator of funds. What evidence should such a person want to see before concluding that a particular intervention is a wise investment? That sort of evidence is what the implementer seeking to justify further funding will need to present.

We begin by setting the conceptual framework within which program costs and outcomes may be understood. We then draw out some of the implications of that general framework for the analysis of early childhood interventions in particular. After reviewing some examples of such analyses, we apply the methodology to an actual case in which a consortium of program funders must decide whether to proceed with an assessment and, if so, what kind of assessment to undertake. The consortium is led by the U.S. Substance Abuse and Mental Health Services Administration and the Casey Family Programs, and the intervention of interest is the *Starting Early Starting Smart* Program.

THE COST AND OUTCOME ANALYSIS FRAMEWORK

Decisionmakers and program implementers just beginning to think about analyzing costs and benefits are often surprised to learn that several analytic avenues are open to them. Which one or ones they choose will have important implications for what they learn and how much they must spend to learn it. Among the choices are these:[1]

- **Cost-benefit analysis** (or benefit-cost analysis) entails comparing a program's benefits to a stakeholder with its costs to that stakeholder. Such a comparison requires putting benefits and costs in comparable terms, and the terms conventionally chosen are dollars. Benefits that cannot be expressed in dollar terms cannot be compared in this manner and are included only in associated qualitative discussion. Cost-benefit analysis seeks to help in deciding whether a program is of value to the stakeholder. Often cost-benefit analysis is conducted from the perspective of society at large.[2]
- **Cost-savings analysis** is restricted to the costs and benefits realized by the government as a whole or a particular funding agency. Only the costs to the government are taken into account, and the benefits are those expressible as dollar savings somewhere in the government. This kind of analysis is used to determine whether a publicly provided program "pays for itself" and is thus justified not only by whatever human services it may render but also on financial terms alone.
- **Cost-effectiveness analysis** determines how much must be spent on a program to produce a particular outcome (or, what is equivalent, how much of a particular type of benefit will result from a given expenditure). While this can be done for multiple outcomes, no attempt is made to sum the complete array of benefits into a single aggregate measure.
- **Cost analysis** alone (no measurement of benefits) can be useful to decisionmakers for a variety of purposes, for example, discovering which factors need to be considered in replicating a program elsewhere or for informing budget projections.

In deciding which avenues to pursue, the decisionmaker or implementer must choose what he or she wishes to learn and consider the funds available for undertaking the analysis. The analyses above are ordered in terms of how much attention must be paid to quantifying outcomes and expressing them in dollar terms (from a lot at the top to none at the bottom). Other variables being equal, the resources and calendar time devoted to the analysis will drop with each successive approach down the list.

---

[1] Terminology in this field has not been standardized, and these terms appear in the literature with a variety of different meanings. We have chosen typical definitions.

[2] Of the four analytic approaches listed here, cost-benefit analysis is subject to the greatest challenges in execution and interpretation. That is because benefits must be denominated in dollars, and that adds another source of uncertainty and potential disagreement over quantities. For some benefits, dollar conversions are not really feasible. Cost-benefit assessments can thus rarely be comprehensive.

As we describe them here, these cost and outcome analysis methods are used only as components within a broader decision support framework that we call *policy analysis* or *policy scorecard analysis* (the latter term derives from the use of a tool called the scorecard).[3] Despite the name, it does not pertain only to high-level public policies but also to decisions made regarding specific strategies and programs. Policy scorecard analysis offers a framework within which to consider multiple benefits, as required in the first two approaches listed above, and multiple costs, as required by all four. Policy scorecard analysis also entails consideration of alternative programs. This is important for benefit and cost analysis. In trying to determine whether the numbers emanating from these analyses support (further) investment in the program, funders will be asking, "compared with investment in what else?" A benefit-cost ratio of 1.5 to one ($1.50 of benefits for every dollar of costs) may not be good enough if an alternative with similar objectives has a ratio of two to one. Decisionmakers will thus be considering a range of alternative interventions or at least a choice between funding the program in question and some default course of action (which could be leaving things as they are).

The results of a policy scorecard analysis can be summarized in a simple tool called a *scorecard*. The scorecard lists benefit and cost categories down the side, together with program design features influencing them, and the alternative courses of action across the top. Thus, each cell in the scorecard gives a particular cost or benefit (or design feature) for a particular program. In identifying the row and column heads and filling in the cells—that is, in conducting the policy scorecard analysis—several guidelines must be kept in mind:

- **Designate which benefits and costs accrue to which stakeholders**. If you say that a program generates more savings than costs, people will want to know, savings to whom? And costs to whom?
- **Define explicitly the period over which the analysis applies**. If the purpose of the analysis is to determine whether a program has a favorable benefit-cost ratio or pays for itself in government savings, it is better to look well into the future. No one period or duration is correct, however. The choice depends on the patience of the decisionmaker in question, with individuals typically having shorter planning horizons than society as a whole. This distinction makes a difference because the costs of early intervention programs typically accrue over a matter of months or a few years, whereas the benefits are often not fully realized until the participating children age into adulthood. Counting such benefits directly entails long-term follow-up of program subjects, though some future benefits can be predicted on the basis of shorter-term trends.
- **Discount future costs and benefits**. Although it is important to count future benefits (and costs), they cannot be counted at full, nominal value. People discount future benefits and costs: getting a $1,000 benefit five years in the future

---

[3] The term *policy analysis* was originally adopted by RAND analysts and others to describe an approach for quantitatively analyzing management problems. Today, the term is used even more broadly to characterize a wide range of quantitative and qualitative approaches to addressing policy issues. Hence, we will employ the more focused term *policy scorecard analysis* for the remainder of this summary.

does not look as attractive as getting it now; having to pay $1,000 five years in the future does not seem as onerous as having to pay it now. A real annual discount rate of 3 percent to 6 percent is typically applied to future benefits and costs.

- **Record cost elements as resource quantities**. Until the figures are added up at the end, costs should be recorded in terms of resource quantities—hours of labor, square footage of rental space, etc.—rather than in dollar terms. Prices for these resources can vary from one site to another, and on-budget dollars in particular do not always reflect total costs. A physician may donate time on the weekends, but from society's point of view, that time is not "free"; perhaps it could have been put to another, more beneficial use.

- **Address uncertainty.** Future benefits and costs cannot often be predicted with great confidence. Where a range of values is plausible, that range should be made explicit in the analysis. Likewise, structural uncertainty (e.g., about possible future changes in laws relevant to a program) should also be considered.

The final step in the cost and outcome analysis is to add up all the benefits (or savings) and add up all the costs and compare them across programs. The four methods listed above offer alternative ways for performing this step. Cost-benefit and cost-savings analysis each provide a single measure of merit for each alternative; the alternative with the greatest merit according to this measure is declared the winner. Cost-effectiveness analysis provides multiple measures of merit. They can be combined into a single measure (e.g., the ratio of effectiveness to cost, if a single effectiveness measure dominates), which will be used in the same way as a cost-benefit or cost-savings measure. Or they can be used to define a different kind of selection rule, one that deems "best" the policy that achieves a specified level of effectiveness at lowest cost (a *constant effectiveness* analysis) or that achieves the greatest effectiveness for a given cost (a *constant cost* analysis).[4]

Comparing costs and benefits may not produce a single "answer" that one program is obviously preferable to another. One program may produce a net benefit to one group of stakeholders, while another benefits a second group. The net benefit of one program may be somewhat higher than that for another, but the uncertainty ranges may overlap so much that the advantage cannot be asserted with high confidence. Some possible change in the institutional environment, e.g., tax reform, could shift benefits and costs enough to change the advantage from one program to another. Such possibilities would not subtract from the value of the cost and outcome analysis. On the contrary, some of the most valuable insights are suggestions for policy changes that reallocate benefits across stakeholder groups so that all of them gain and thus have no incentive to block a program.

In most studies, the majority of the analytical effort will come from learning about the domain, structuring the models of how the intervention works, collecting and cleaning data, etc. In short, filling in the scorecard is challenging. Given that groundwork, computing the summary evaluation metrics is straightforward, whether that metric is a benefit-cost or a cost-effectiveness ratio.

---

[4] The latter is sometimes called a *constant budget analysis*, but this is only appropriate if all the costs appear in the budget of the agency making the decision. In many programs, costs may be distributed across many stakeholders. They will not all appear in any single party's budget.

Hence, instead of suggesting that one must choose to implement one of these four approaches, it is more accurate to say that one must choose whether or not to conduct a careful, quantitative summation of the effects of the program. If the answer is yes, then there follows a choice of how one is going to present the results of that analysis to decisionmakers, as a benefit-cost ratio, cost-effectiveness ratio, and so on, or some combination thereof.

It is thus important to keep cost-benefit analysis, cost-savings analysis, and other forms of cost and outcome analysis in their place. In any decision, some factors can be resolved only through a decisionmaker's values and subjective judgment or through negotiation among stakeholders. Likewise, the public quantifying of decision factors may occasionally be problematic (e.g., when an auto manufacturer compares the cost of a safety improvement with the dollar-equivalent benefit of the lives that could be saved by that design change). Nevertheless, these methods can provide valuable input to choosing among different programs, demonstrating a program's worth, improving programs, and replicating them.

APPLYING THE FRAMEWORK TO EARLY CHILDHOOD INTERVENTIONS

Early intervention programs attempt to improve child health and development by providing young children and their families various social services and supports. Such programs can have effects in four domains: emotional and cognitive development, education, economic well-being (in terms of public assistance, income, and crime), and health. Specific examples of possible benefits within each of these categories are given in Table S.1. Which benefits are measured depends on the purpose of the analysis. Cost-benefit and cost-savings analyses typically seek a comprehensive accounting of the benefits to society or to government (respectively), although many benefits are difficult to express in dollar terms and therefore cannot be aggregated in the cost-benefit assessment. While cost-effectiveness analysis can in principle be performed for any outcome, it is often the case in practice that a single benefit or a narrow set receives most of the attention. A full analysis of the benefits of an early intervention program should include collection of data on as many potential benefits as the analyst's resources permit.

Note that early childhood interventions can benefit parents and other caregivers while simultaneously helping children. It is important to measure benefits to caregivers, because these are often realized over much shorter time periods than are those accruing to children. Ignoring these benefits means underestimating a program's benefit-cost ratio or its potential net savings to government, particularly over the short term—and for some analyses, it will only be feasible to make short-term measurements.

Any analysis of benefits of a program under way *must* include a comparison group. This is a group of children and caregivers not enrolled in the program but similar in

Table S.1

Early Childhood Intervention Program Benefit Domains and Illustrative Measures

	Illustrative Measures for:	
Benefit Domain	Child	Parent/Caregiver
Emotional and cognitive development	Socioemotional and behavior scores IQ test scores Teacher's ratings	Quality of parent-child relationship Quality of home environment
Education	Achievement test scores Grades *Grade progression (repetition)* *Participation in special education* *Educational attainment*	*Educational attainment*
Public assistance receipt, income, crime	*Receipt of public assistance* *Employment* *Earnings/income* *Criminal activity* *Contact with criminal justice system*	*Receipt of public assistance* *Employment* *Earnings/income* *Criminal activity* *Contact with criminal justice system*
Health	Physical and mental health status *Child abuse and neglect* Substance abuse Fertility control *Emergency room visits* *Other health care use*	Physical and mental health status Family violence Substance abuse Fertility control

NOTE: *Italics* indicate measures more easily expressed in dollar terms.

as many ways as possible to the program participants and whose progress along the various benefit measures is tracked.[5] Children in particular have a tendency to improve along various measures of development as they grow. Evaluators must take care to ensure that the program benefits they measure are net of what would have occurred naturally or what children would realize anyway from outside influences without the program. Measurements of the comparison group provide estimates of benefits that would have accrued in the program's absence.

Data on progress along benefit measures can be collected by survey questionnaires, tests, or other means of direct interaction with the children and their caregivers. For some benefit types (e.g., reductions in involvement with the criminal justice system), administrative data may be available. When only a few years of data collection are feasible, a glimpse into the future can be obtained through mathematical models that can predict future criminal activity or future earnings on the basis of childhood information. (This cannot of course be done with confidence for any given child, but results obtained for a group of children may be sufficiently reliable for the purpose.)

---

[5] Ideally, one should randomly assign children and caregivers to program participation versus the comparison group. This ensures that the participation and comparison groups are (statistically) identical in both measured and unmeasured characteristics. When the comparison group is selected by random assignment, it is often called a control group. When random assignment is not feasible or desirable, a comparison group can still be chosen, by identifying children and caregivers who are similar in various measured ways to the program participants.

As with benefits, the cost elements to be included in an analysis depend on its purpose. For example, costs that accrue to society but not to a funding agency are included in a societal cost-benefit analysis but omitted from a cost-savings analysis. Regardless of the analysis to be performed, program costs must be estimated as net of those accrued by comparison group children for similar services. For example, if an intervention is intended to increase prenatal care, the analysis should include only the resources devoted to the visits and services received by program participants in excess of what they would have received anyway (i.e., in excess of those received by the comparison group).

Estimation of costs should follow the general guideline given above regarding the need to estimate resource quantities instead of dollars and to account for "opportunity" costs and other off-budget resource expenditures. Costs borne by participants should also be included, as well as costs borne by other agencies or service providers. Collecting cost data for the same set of service providers for both the treatment and control groups allows the analyst to detect both cost shifting (e.g., from one payor to another) and cost offsets (e.g., reduced utilization of services in one area as a result of increased service use in another). It may also be useful to distinguish between the fixed costs of implementing a program that are not dependent on the number of children served and the variable costs that are. The split between fixed and variable costs will influence the calculation of benefit-cost ratios, net savings, and cost-effectiveness ratios for programs when scaled up to serve larger numbers of children.

SOME ILLUSTRATIVE ANALYSES

Given the challenges and requirements outlined so far, it should not be surprising that not many scientifically sound cost-benefit and cost-savings analyses of early childhood intervention programs with long-term follow-ups have been conducted. Among those recently analyzed or reanalyzed are the following:

- The **Perry Preschool** program provided center-based classes and teacher home visits for one or two school years to 58 children ages three or four in Ypsilanti, Michigan, from 1962 to 1967. Benefits were tracked for both the participants and the comparison group (65 children) through age 27. Benefits included better school performance, higher employment, less welfare dependence, and lower involvement in criminal activity on the part of participants. The most recent cost-benefit assessment evaluates benefits expressible in monetary terms at $50,000 per child, half of that in the form of savings to government, versus a program cost of $12,000 per child (see Figure S.1).
- In the **Prenatal/Early Infancy Project** (PEIP) in Elmira, New York, nurses started visiting mothers when they were pregnant and continued until their child was age two. The objective was to improve pregnancy outcomes and parenting skills

and link the mother with social services. Between 1978 and 1980 the program reached 116 first-time mothers. They and another 184 in the control group have been followed through age 15 of the first-born child. Benefits for the mothers included better pregnancy behaviors and less child abuse in the short term and lower welfare participation and criminal behavior in the long term. The children benefited as well in several domains. For the higher-risk portion of the sample (unmarried mothers with low socioeconomic status), benefits amounted to almost $31,000 per mother-child pair, with almost half of that in the form of a reduction in welfare received by the mother. For the lower-risk portion of the sample, however, benefits came to only $6,700. Program costs were about $6,100.

- The **Chicago Child-Parent Centers** have promoted reading and language skills, provided health and social services, and promoted parent involvement for children in preschool through third grade. A cohort of 989 children completing kindergarten in 1986 was tracked to age 20 and compared with a no-preschool group of 550 children. The program resulted in long-lasting educational-achievement benefits. Higher between-grade promotion rates, reduced special-education use, increased earnings expected as a result of better educational performance, and lower involvement with the juvenile justice system translated into about $35,000 in benefits per program participant. The program cost nearly $10,000 per participant.

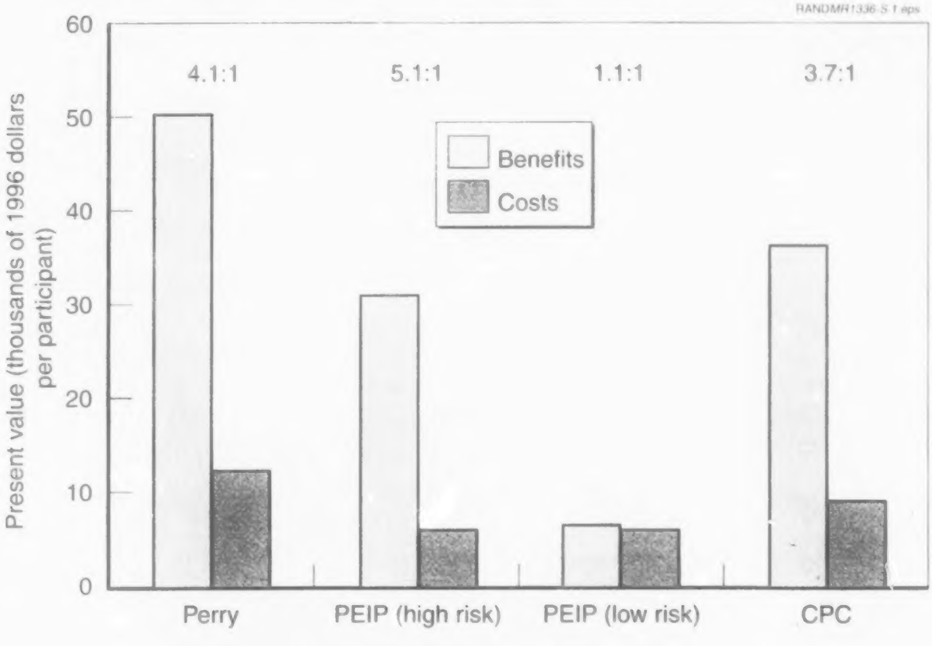

Figure S.1—Some Early Childhood Interventions Have Been Shown to Have High Benefit-Cost Ratios

These analyses demonstrate that early childhood interventions can generate savings to government and benefits to society that exceed program costs. Indeed, for most of the samples reported above, benefits were a multiple of costs, and all of these programs resulted in benefits that could not be translated into costs and were thus omitted. Therefore, decisionmakers and implementers thinking about performing analyses of costs and benefits should not give up merely because they don't see how some of a program's principal benefits can be converted to dollar terms.

Two further lessons for cost-benefit analysis may be drawn from these examples. First, many important benefits can only be captured through an extended time horizon. The savings from Perry Preschool, for example, did not accumulate to match the level of program costs until the participants were 20 years old. Some of these benefits can be predicted on the basis of shorter trends, but not all can, and confidence in predicted results increases as follow-up periods lengthen.

Second, programs can be beneficial to caregivers as well as to children. In fact, when time is lacking for lengthy follow-ups or when they are not feasible, measuring benefits to caregivers can result in early favorable benefit-cost ratios and net savings. The Elmira program was the only one of those summarized that measured caregiver benefits, and, in that case, savings sufficient to balance costs were tallied within two years of the end of program services.

FRAMING A POLICY SCORECARD ANALYSIS FOR A SPECIFIC PROGRAM

The *Starting Early Starting Smart* (SESS) program is intended to test the effectiveness of integrating mental health services and substance abuse prevention and treatment into early childhood education or primary health care for children from birth to age seven. The program is under way at 12 sites nationwide, seven using the early childhood (EC) education model and five using the primary care (PC) paradigm. (See the appendix of the full report [forthcoming] for a description of each state.) Most of the sites serve between 100 and 300 children, and comparison groups average out to similar numbers.

By "effectiveness," the program means increased access to, use of, and satisfaction with behavioral health services and increased social, emotional, and cognitive functioning on the part of served children. Data on these benefit measures are being collected over an 18-month follow-up period at intervals that average six months (PC sites) or nine months (EC sites). No cost data are being gathered in this first phase of the program, but a second phase is being planned, and part of that planning is to assess the feasibility of cost and outcome analysis.

SESS program implementers are wise to take cost and benefit evaluation issues into account in the planning stage. Too often, evaluation is considered only after program design has been finalized along lines that preclude sound cost and benefit assessment. SESS's Phase I design raises issues that need to be resolved for Phase II if cost and outcome analyses are to be possible. One issue, for example, is that some sites did not use random assignment (primarily EC sites), which raises concerns about the validity of the treatment group versus comparison group difference as a measure of the true effects of the program. Future demonstration sites should aim for random assignment if at all possible. Another concern is that a few sites are experiencing relatively high dropout rates, which could bias benefit estimates if those who are lost to follow-up are different from those who remain in the study and if they differ in important ways that cannot be observed. Obtaining a consistently high follow-up rate across sites would need to be a priority in Phase II. Also, Phase I has been characterized by between-site variations in services. This is problematic from an evaluation standpoint for a couple of reasons: It complicates interpretation of results, and it complicates the design of comparison groups.

The design of comparison groups for SESS offers lessons for other programs. Because SESS attempts to integrate behavioral health services into existing early childhood and primary-care settings, only the benefits of the new, integrated services plus increases in the "dosages" of existing services may be credited to SESS, not the full benefits realized from participation in the early childhood program and primary care. Similarly, only the costs associated with these incremental activities should be considered. Therefore, the comparison groups must be designed to isolate the SESS effects by including everything except SESS. The appropriate comparison groups for this evaluation would consist of children involved in early childhood and primary-care programs without the integrated SESS services, not children receiving no services at all.

In the policy analysis scorecard, then, the columns would correspond to the early childhood program without SESS, primary-care program without SESS, and then the integrated EC plus SESS and PC plus SESS interventions, along with whatever variants are retained. The rows would be the program descriptors and cost and benefit categories. The program features reported would be those having implications for costs or benefits, e.g., population served, eligibility criteria, age of children at enrollment, qualifications of program personnel, types and "dosages" of services rendered, transportation provisions, and so on. In future demonstrations, this information can be collected through site visits and other mechanisms currently being used in the evaluation of Phase I.

Cost estimates would begin with the cost of serving one child (or child's caregiver) in terms of labor hours expended with the child and in preparing for the session and in terms of materials consumed. These would then be multiplied by dosage per child and number of children served. Fixed costs unrelated to number of children served, such as space rental, would then be identified. Multiplication by unit costs to convert to dollars would be done last. Ultimately, the cost information should be as comprehensive as possible and comparable across demonstration sites.

Benefit measures now being collected for SESS include information on child problem behavior and social skills, child cognitive development, parent-child interaction, caregiver stress and negative or positive behaviors, caregiver mental health problems, caregiver education and employment, and home environment. As discussed above, the emphasis on both child and caregiver benefits will be important to making the short-run benefit tally as complete as possible. Almost all of these measures, however, are within the domain of emotional and cognitive development and are not easily expressed in dollar terms. This makes a formal cost-benefit or cost-savings analysis problematic in that only a limited set of outcomes might possibly be valued in dollar terms to be compared with program costs. Unless the program impact for those outcomes valued in dollar terms is very large and favorable, so that sizable dollar benefits are generated, a cost-benefit analysis would be unlikely to show a favorable outcome for the SESS program based on the information available after two years.

While not the program's main intent, other benefits could result from it. Some of these benefits, in such areas as physical health, labor market outcomes, and involvement with the criminal justice system, could be more easily expressed in dollar terms than those now being measured. These outcomes could be collected for parents or caregivers in the short term, and with longer-term follow-up, for the participating children. If behavioral changes are large in these areas as a result of the SESS intervention, they can produce sizable dollar benefits that, even when discounted, will be a large offset to the costs of the program. This is especially relevant for changes in parental behavior that can be measured even in the short run. Improvements of adult economic and health outcomes have been demonstrated to produce substantial short-run benefits in other early childhood programs.

Costs and outcomes would be measured for both the participant and comparison groups, with the difference between the two constituting the incremental cost and benefits from implementing SESS. To compare the present values of all costs and benefits, it will be important to predict how they will accrue over time. Costs and benefits should also be categorized according to which groups incur them. It will be of interest, for example, to know how much the intervention costs and benefits participants, the agency implementing the program, other agencies, and society as a whole.

Taking all these steps would be sufficient to support as full a cost-benefit or cost-savings analysis as is likely to be feasible given the current state of the art. If SESS decisionmakers wish to be able to say something about the value the program returns to society relative to its costs, the preceding array of evaluation tasks and program design modifications would be required. If they decide it is enough to be able to say how much the program saves the government relative to what it costs, then some elements—costs to participants or losses to crime victims, for example—can be omitted. The overall level of effort required, however, is not likely to change very much.

If SESS funders or implementers would like instead to focus on one or a few prominent measures of effectiveness to compare the different SESS variants with each other, a cost-effectiveness analysis should be sufficient. By collecting cost data, along with data on that one or those few benefits, it would be possible to say, for example, how much child problem behavior decreased (relative to no SESS) per thousand dollars spent on SESS plus EC or SESS plus PC. No conversion of the benefit to dollar terms would be necessary.

Finally, if the purpose was to find out how much program modifications or proliferation of sites would cost, no benefit data would be necessary at all. Clearly, program decisionmakers may have to make trade-offs between what they might like to achieve and how much of a resource commitment they are willing or able to make.

CONCLUSIONS

The recommendations we offer specific to the SESS program may be framed as a set of more-general guidelines for decisionmakers considering cost and outcome analysis of an early childhood intervention program. In particular, among the recommendations that can be applied more broadly are the following:

- Regarding the design of a program evaluation and cost and outcome analysis:
    - Specify the explicit goals of the cost and outcome analysis to guide the scope of cost and benefit data collection and analysis.
    - Identify comparison groups and track the same cost and outcome measures for both comparison and participant groups. If possible, use random assignment to define comparison groups to provide a more valid test of intervention program effects.
    - To minimize attrition in a longitudinal study, devote resources to retaining study subjects.
    - Collect information on program features through site visits and other mechanisms to accurately characterize features of the intervention models as they are implemented and to ensure fidelity to the program model.
- Regarding the collection and analysis of cost data:
    - Collect cost information for both treatment and comparison groups at each site where the intervention program is implemented.
    - Ensure that the cost information is as comprehensive as possible: Costs borne by various parties should be differentiated, the period in which costs are incurred should be identified, and direct and indirect costs, fixed and variable costs, and goods and services provided in-kind should be measured.
    - Plan for proper training and technical support of implementation sites and any cross-site data collection organizations to ensure uniformity in the collection of cost data. Collect information on the cost of data collection, training and support, and the related analyses of the data.
- Regarding the collection and analysis of outcome data:
    - If cost-benefit or cost-savings analysis is the goal, include in the outcome data information for parents and other caregivers in the short term and long term and for children in the long term in those domains with outcomes that can be readily evaluated in terms of dollars and can produce large dollar benefits. The choice of specific outcome measures should be guided by findings from related evaluation studies whenever possible.

- Obtain information from participants that facilitates collection of administrative data and allows effective tracking of individuals to increase response rates at later follow-ups.
- Where possible, collect complete histories using retrospective survey questions or administrative data for outcomes that may generate a continuous flow of dollar benefits (e.g., labor market outcomes, social welfare program use, use of costly health or education services).
- When supported by other empirical evidence, project future benefits based on observed outcomes. Consider additional method development that would permit such forecasts for a broader range of outcomes.

While we believe these principles are quite general, ultimately these recommendations should be viewed as guidelines that may need to be tailored to the specific circumstances of a given intervention program and its evaluation design. In the end, the objectives of a program's decisionmakers will dictate the shape of the analysis.

The general policy scorecard analysis tools considered in this report, and those specific to cost and outcome analysis, have great promise for improving decisionmaking with respect to such investment programs as the early childhood interventions represented by SESS and its counterparts. When used with skill and judgment, the application of these methods to other programs, such as SESS, will further broaden our base of knowledge regarding the value of these investments and aid decisionmakers in their choice among program alternatives.

STARTING EARLY STARTING SMART GRANT SITES

The SESS program is an initiative of the Office on Early Childhood, Substance Abuse, and Mental Health Services Administration (SAMHSA) and the Casey Family Programs, along with other federal sponsors. Patricia Salomon, Director of the Office of Early Childhood at SAMHSA, oversees the SESS program along with project officers Michele Basen, Velva Spriggs, and Jocelyn Whitfield, and staff Shakeh Kaftarian. At the Casey Family Programs, the partnership is overseen by Jean McIntosh and Barbara Kelly-Duncan, along with project officers Eileen O'Brien and Peter Pecora.

The SESS program currently operates in 12 sites across the U.S. Table A.1 lists each of the study sites and the associated principal investigator, project director, and local researcher, first for the primary care (PC) sites and then for the early childhood (EC) sites.[1] Information about the Data Coordinating Center is also provided in Table A.1. A brief description of the program at each site is provided in the appendix to the companion report. Further information about the SESS program is provided in Appendix B and Appendix C and is available from the Casey Family Programs (www.casey.org/projects.htm#sess) and SAMHSA (www.samhsa.gov).

Table A.1

SESS Grant Sites

Study Site	Principal Investigator	Project Director	Local Researcher
	Data Coordinating Center		
EMT Associates, Inc., Folsom, Calif., (615) 595-7658	Joel Phillips	J. Fred Springer, Ph.D.	J. Fred Springer, Ph.D.
	Primary Care Sites		
Boston Medical Center, Boston, Mass., (617) 414-7433	Carol Seval, R.N., L.M.H.C.	Carol Seval, R.N., L.M.H.C.	Ruth Rose-Jacobs, Sc.D.
The Casey Family Partners, Spokane, Wash., (509) 473-4810	Christopher Blodgett, Ph.D.	Mary Ann Murphy, M.S.	Christopher Blodgett, Ph.D.

---

[1] One of the original SESS sites was unable to continue with the study but was an important contributor to the original design and implementation of the project.

Table A.1—continued

Study Site	Principal Investigator	Project Director	Local Researcher
University of Miami, Miami, Fla., (305) 243-2030	Connie E. Morrow, Ph.D.	K. Lori Hanson, Ph.D.	Emmalee S. Bandstra, M.D. April L. Vogel, Ph.D.
University of Missouri, Columbia, Mo., (573) 884-2029	Carol J. Evans, Ph.D.	Robyn S. Boustead, M.P.A.	Carol J. Evans, Ph.D.
University of New Mexico, Albuquerque, N.M., (505) 272-3469	Andy Hsi, M.D., M.P.H.	Bebeann Bouchard, M.Ed.	Richard Boyle, Ph.D.
	Early Childhood Sites		
Asian American Recovery Services, Inc., San Francisco, Calif., (415) 541-9285, ext. 227	Davis Y. Ja, Ph.D.	Anne Morris, Ph.D.	Anne Morris, Ph.D.
Child Development Inc., Russellville, Ark., (501) 968-6493	JoAnn Williams, M.Ed.	Carol Amundson Lee, M.A., L.P.C.	Mark C. Edwards, Ph.D.
Children's National Medical Center, Washington, D.C., (202) 884-3106	Jill G. Joseph, M.D., Ph.D.	Amy Lewin, Psy.D.	Michelle J. C. New, Ph.D.
Johns Hopkins University, Baltimore, Md., (410) 955-3989	Philip J. Leaf, Ph.D.	Jocelyn Turner-Musa, Ph.D.	Philip J. Leaf, Ph.D.
Division of Child and Family Services, Las Vegas, Nev., (702) 486-6147	Christa R. Peterson, Ph.D.	Laurel Swetnam, M.A., M.S.	Margaret P. Freese, Ph.D., M.P.H.
The Tulalip Tribes Beda?chelh, Marysville, Wash., (360) 651-3282	Linda L. Jones, B.A.	Linda L. Jones, B.A.	Claudia Long, Ph.D.
The Women's Treatment Center, Chicago, Ill., (773) 373-8670, ext. 302	Jewell Oates, Ph.D.	Dianne Stansberry, B.A., C.S.A.D.P.	Victor J. Bernstein, Ph.D.

SESS PROGRAM ACKNOWLEDGMENTS

The families and grantees of *Starting Early Starting Smart* (SESS) would like to acknowledge:

Nelba Chavez, Ph.D. Administrator, SAMHSA Rockville, Maryland	Ruth Massinga, M.S. President and CEO, Casey Family Programs Seattle, Washington

along with the Casey Board of Trustees and the three SAMHSA Centers—Center for Substance Abuse Prevention, Center for Substance Abuse Treatment, and Center for Mental Health Services—for their vision and commitment to reaching families with very young children affected by environments of substance abuse and mental disorders. Without their innovative public-private partnership and unprecedented support, this initiative would have been impossible.

We further acknowledge the early guidance and program development from Stephania O'Neill, M.S.W., Rose Kittrell, M.S.W., Hildy (Hjermstad) Ayers, M.S.W., Karol Kumpfer, Ph.D., Sue Martone, M.P.A., and Jeanne DiLoreto, M.S.

Many thanks to the SAMHSA-Casey team for their tenacious efforts and unprecedented collaboration:

Joe Autry, M.D. Acting Administrator SAMHSA	Jean McIntosh, M.S.W. Executive Vice President Casey Strategic Planning and Development
Pat Salomon, M.D.	Barbara Kelley Duncan, M.Ed.
Michele Basen, M.P.A.	Peter Pecora, Ph.D.
Velva Springs, M.S.W.	Eileen O'Brien, Ph.D.
Jocelyn Whitfield, M.A.	

MISSION STATEMENTS OF THE NATIONAL COLLABORATORS

SUBSTANCE ABUSE AND MENTAL HEALTH SERVICES ADMINISTRATION (SAMHSA)

SAMHSA's mission within the nation's health system is to improve the quality and availability of prevention, treatment, and rehabilitation services to reduce illness, death, disability, and cost to society resulting from substance abuse and mental illness.

SAMHSA's mission is accomplished in partnership with all concerned with substance abuse and mental illness. SAMHSA exercises leadership in

- eliminating the stigma that impedes prevention, treatment, and rehabilitation services for individuals with substance abuse;
- developing, synthesizing, and disseminating knowledge and information to improve prevention, treatment, rehabilitation services, and improving the organization, financing, and delivery of these services;
- providing strategic funding to increase the effectiveness and availability of services;
- promoting effective prevention, treatment, and rehabilitation policies and services;
- developing and promoting quality standards for service delivery;
- developing and promoting models and strategies for training and education;
- developing and promoting useful and efficient data collection and evaluation systems; and
- promoting public and private policies to finance prevention, treatment, and rehabilitation services so that they are available and accessible.

For more information, visit SAMHSA's Web site at www.SAMHSA.gov.

CASEY FAMILY PROGRAMS

The mission of Casey Family Programs is to support families, youth, and children in reaching their full potential. Casey provides an array of permanency planning, prevention, and transition services, such as long-term family foster care, adoption, kinship care, job training, and scholarships.

The program aims to improve public and private services for children, youth, and families impacted by the child welfare system, through advocacy efforts, national and local community partnerships, and by serving as a center for information and learning about children in need of permanent family connections.

Casey Family Programs is a Seattle-based private operating foundation, established by Jim Casey, founder of United Parcel Service (UPS), in 1966. The program has 29 offices in 14 states and Washington, D.C. For more information, visit our Web site at www.casey.org.

REFERENCES

Adler, M. D., and Posner, E. A. (2000). "Introduction," to Cost-benefit analysis: legal, economic, and philosophical perspectives, a conference sponsored by the John M. Olin Foundation and University of Chicago Law School, *Journal of Legal Studies, 29* (2, part 2), 837–842.

Barnett, S. W. (1993). Benefit-cost analysis of preschool education: Findings from a 25-year follow-up. *American Journal of Orthopsychiatry, 63* (4), 500–508.

_____ (1995). Long-term effects of early childhood programs on cognitive and school outcomes. *The Future of the Children, 5*, 25–50.

Cannon, J. S., Karoly, L. A., & Kilburn, M. R. (2001). *Directions for cost and outcome analysis of* Starting Early Starting Smart*: Summary of a cost expert meeting.* Santa Monica, CA: RAND, CF-161-TCFP.

Caulkins, J. P., Rydell, C. P., Everingham, S. S., Chiesa, J., & Bushway, S. (1999). *An ounce of prevention, a pound of uncertainty: The cost-effectiveness of school-based drug prevention programs.* Santa Monica, CA: RAND, MR-923-RWJ.

Frank, R. H. (2000). Why is cost-benefit analysis so controversial? *The Journal of Legal Studies, 29* (2), Part 2, 913–930.

Guralnick, M. J. (Ed.) (1997). *Effectiveness of early intervention.* Baltimore: Paul Brookes Publishing.

Hargreaves, W. A., Shumway, M., Hu, T.-W., & Cuffel, B. (1998). *Cost-outcome methods for mental health.* San Diego: Academic Press.

Karoly, L. A., Greenwood, P. W., Everingham, S. S., Houbé, J., Kilburn, M. R., Rydell, C. P., Sanders, M., & Chiesa, J. (1998). *Investing in our children: What we know and don't know about the costs and benefits of early childhood interventions.* Santa Monica, CA: RAND, MR-898-TCWF.

Kitzman, H., Olds, D. L., Henderson, C. R., et al. (1997). Effect of prenatal and infancy home visitation by nurses on pregnancy outcomes, childhood injuries, and repeated childbearing: A randomized controlled trial. *Journal of the American Medical Association, 278* (8), 644–652.

Olds, D. L., Eckenrode, J. Jr., Henderson, C. R., et al. (1997). Long-term effects of home visitation on maternal life course, child abuse and neglect, and children's arrests: Fifteen-year follow-up of a randomized trial. *Journal of the American Medical Association, 278* (8), 637–643.

Posner, R. A. (2000). Cost-benefit analysis: Definition, justification, and comment on conference papers. *The Journal of Legal Studies, 29* (2), Part 2, 1153–1177.

Reynolds, A. J. (2000). *Success in early intervention: The Chicago Child-Parent Centers.* Lincoln, NE: University of Nebraska Press.

Richardson, H. S. (2000). The stupidity of the cost-benefit standard. *The Journal of Legal Studies, 29* (2, part 2), 971–1003.

Schweinhart, L. J., Barnes, H. V., & Weikart, D. P. (1993). *Significant benefits: The High/Scope Perry Preschool Study through age 27.* Ypsilanti, MI: High/Scope Educational Research Foundation, Monographs of the High/Scope Educational Research Foundation, No. 10.

Sen, A. (2000). The discipline of cost-benefit analysis. *The Journal of Legal Studies, 29* (2, part 2), 931–952.

The SESS Sites

Miami's Families: Starting Early Starting Smart	Florida
Raising Infants in Secure Environments	Massachusetts
Healthy Foundations for Families	Missouri
Starting Early to Link Enhanced Comprehensive Treatment Teams	New Mexico
Casey Family Partners	Washington
National Association for Families and Addiction Research and Education	Illinois*
Child Development, Inc.	Arkansas
Asian American Recovery Services, Inc.	California
Locally Integrated Services in Head Start	Washington, D.C.
Starting Early Starting Smart Head Start Collaboration Project	Illinois
Baltimore BETTER Family and Community Partnership	Maryland
New Wish	Nevada
Beda?chelh Tulalip Tribes Early Intervention in Tribal and Mainstream Communities	Washington
Evaluation, Management and Training, Inc.**	California

*One of the original SESS sites was unable to continue with the study, but it was an important contributor to the original design and implementation of this project. Our thanks to Dr. Linda Randolph and Dr. Ira Chasnoff.

**Data Coordinating Center

BEST COPY AVAILABLE

Please feel free to be a "copy cat" by making all the copies you want of the entire document; or if sections are copied, please provide the full citation to the report.

For further information
on alcohol, tobacco,
and illicit drugs
call 1-800-729-6686
301-468-2600, or TDD
1-800-487-4889
Or visit the World
Wide Web at
www.health.org

SAMHSA

CASEY
FAMILY
PROGRAMS
FOUNDATIONS FOR THE FUTURE

STARTING
EARLY
STARTING
SMART

Healthy Minds, Healthy
Behavior: Promoting
Life—Right from the Start

BEST COPY AVAILABLE

CMHS
Center for Mental Health Services
SAMHSA

CSAT
Center for Substance
Abuse Treatment
SAMHSA

CSAP Center for
Substance Abuse
Prevention
Substance Abuse and Mental
Health Services Administration